

# nature

THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

## FICKLE FATES

Physics of  
asymmetric  
endosomes  
during asymmetric  
cell fate assignment

PAGES 196 & 280

FETAL RESEARCH

### COURTING CONTROVERSY

The real story of fetal  
tissue research

PAGE 178

SUSTAINABILITY

### CHARTING THE DATA

How evidence mapping  
unearths the facts that count

PAGE 185

METHODS

### NEW LIGHT ON CELL BIOLOGY

Optogenetics revolutionized  
neuroscience — what next?

PAGE 291

NATURE.COM/NATURE

10 December 2015 £10

Vol. 528, No. 7581



9 770028 083095



# THIS WEEK

## EDITORIALS

**FUTURE-PROOF** Short-term solutions to long-term issues **p.164**

**WORLD VIEW** Business must pull up its sustainable socks **p.165**



**SEA ME?** Cuttlefish switch off to evade marine predators **p.166**

## Fetal tissue research under threat

*The US Senate has just voted to defund one of the providers of aborted fetal tissue for research. Such research is too valuable to become embroiled in the bitter abortion debate.*

When a journalist invites scientists to discuss their work in the pages of *Nature*, it is rare to encounter a resounding silence. But that was the case when our reporter reached out to biologists in the United States this autumn to ask about the value and applications of their research with human fetal tissue. Just two of the 18 scientists we contacted were willing to go on the record with details of their work.

The reticence is understandable. A hostile political climate surrounds this research in the United States, where the release in July of covertly filmed videos ignited a firestorm of controversy.

Made by anti-abortion campaigners, posing as executives of a fictional biological-supply company, the videos showed senior physicians from the Planned Parenthood Federation of America frankly discussing their supply of legally aborted human fetal tissue for research.

The videos insinuated that the non-profit health-care provider was breaking the law by supplying the fetal tissue to biological-products companies for financial gain. But despite the numerous leading questions, the videos show no law-breaking. In exchange for the fetal tissue, the organization received only legally allowable costs: less than US\$100 for each specimen, at 1% of its 700 clinics. If Planned Parenthood, which mainly provides contraception, cancer screening and other important health care, was seeking to get rich, it chose a strange way to do so.

That has not stopped Republican politicians from seizing on the videos to make repeated, inaccurate and inflammatory accusations. Presidential hopeful Marco Rubio, a US senator from Florida, charged, with utterly no evidence, that the collection of fetal tissue has “created an incentive for people to be pushed into abortions so that those tissues can be harvested and sold for a profit”. Ted Cruz, a US senator from Texas who is also contending for the Republican presidential nomination, declared that Planned Parenthood is “an ongoing criminal enterprise”.

It is not surprising then that, since July, even the small number of Planned Parenthood clinics supplying fetal tissue has dwindled. Or that when an unhinged gunman launched a murderous rampage last month, he chose a Planned Parenthood clinic in Colorado as a target.

Nor is it surprising that US scientists who use fetal tissue are choosing to stay silent about the value of their work rather than to defend it publicly and face the real possibility of physical attack. (One scientist told *The New York Times* that in response to threats against him his institution had posted a guard outside his lab.) The two US-based biologists who did speak to *Nature* should be applauded for their courage.

As the News Feature on page 178 shows, research that uses fetal tissue is worth defending. And there are ways in which the scientific community can rally round without putting individuals at risk. Admirably, the Association of American Medical Colleges (AAMC) is showing the way. The AAMC released a statement last week signed by 58 academic medical centres, scientific societies and allied groups.

The statement outlined the medical advances that have been made possible by fetal tissue, and described the value of its current applications in areas such as developmental biology and research on infectious diseases. The authors wrote of their “grave concerns” about the numerous legislative proposals now in play in the US Congress and in a dozen states — proposals that would restrict or prohibit fetal tissue research. They warned eloquently that the proposed laws “would limit new research on vaccines not yet developed, for treatments not yet discovered, for causes of diseases not yet understood”. *Nature* shares the authors’ grave concerns, and joins the AAMC in calling on US lawmakers to reject proposals that restrict access to fetal tissue.

**“It is time for a de-escalation of the rhetoric and the creation of a space for calm and rational discourse.”**

The current episode is a reflection of a larger politics of division that has taken hold in the United States, and which has worsened alarmingly in recent months. It is time for a de-escalation of the rhetoric and the creation of a space for calm and rational discourse.

In the case at hand, that could begin with greater separation of the issues of fetal tissue research and abortion. Clearly, there is fair, honest and understandable disagreement on the morality of the latter. In a democracy, opponents of abortion are free to do their best within the law to change the law. But nobody benefits when they target by proxy an activity that is tangential to the act that they abhor and that is doing a great deal to advance our understanding of health and disease. ■

## Stem the tide

*Japan has introduced an unproven system to make patients pay for clinical trials.*

Japan has been working feverishly to stay at the cutting edge of research and clinical applications in regenerative medicine. It has invested billions of yen in induced pluripotent stem (iPS) cells — made by reprogramming an individual’s adult cells so that they can develop into any body tissue — and has overhauled its drug regulations to create a fast track to bring regenerative therapies to market.

The strategy is working, up to a point — in September, the first treatments were approved under the new law. According to bullish regenerative-medicine firms in Japan, the scheme is the fastest way to meet patients’ needs. Without it, they argue, treatments get bogged down in phased clinical trials that can take several years and cost hundreds



of millions of dollars. But it is not clear whether the acceleration will benefit patients or help Japan's overburdened national health system.

One of the approved treatments, HeartSheet, is made of skeletal-muscle stem cells that are taken from a patient's thigh and grown in the lab. The sheet, made by the company Terumo, is then applied to the hearts of people who have severe cardiac failure. Japan's health ministry gave "conditional approval" for clinical use of the treatment after the company carried out a phase II trial, which hinted at its safety and efficacy in seven patients (Y. Sawa *et al. Circ. J.* **79**, 991–999; 2015).

The company can market and sell the treatment. The approval comes with the condition that, within 5 years, Terumo must provide data from at least 60 patients treated with HeartSheet and 120 controls to show that the treatment is effective. Officials at the Pharmaceuticals and Medical Devices Agency, which approves new treatments, say that the examination of these data will be just as strict as it would be for a conventional phase III clinical trial.

Such approvals feed two Japanese obsessions. First, they allow Japan to be at the forefront of regenerative medicine, something that it has pursued doggedly since iPS cells — which would go on to win one of the country's scientists a Nobel prize — became a national project. Second, Japan is determined to find new engines of economic growth, because it has enjoyed few successes in biotechnology so far.

Biotech firms around the world are excited about the approval, too. Stories of commercialization are a welcome counterpoint to the narrative of failure. California biotech firm Geron, once a trailblazer in regenerative medicine, has given up on embryonic stem-cell therapies and, just this year, Masayo Takahashi of the RIKEN Center for Developmental Biology in Kobe decided to halt her trial of iPS-cell-derived retinal grafts to treat age-related macular degeneration.

Patients are willing to pay, and pay dearly: the HeartSheet treatment costs nearly ¥15 million (US\$122,000). Last month, the health ministry added it to the procedures covered by national health insurance, which will help. But patients still pay 10–30% of the cost for a

drug that is not known to be effective. As they do so, they basically subsidize the company's clinical trial.

Japan has turned the drug-discovery model on its head. Usually, the investment — and thus the risk — is borne by drug companies, because they stand to gain in the long run. Now the risk is being outsourced. By the time it is clear whether a treatment works or not, the companies will have already made revenue from it.

The government argues that its system will encourage firms to bring to market regenerative-medicine treatments that might work. They

will, at least, work well enough to make it past small initial trials. Many drugs do that, and then most of them fail at phase III.

Biotech companies in other countries are keen on the idea and have pushed their own regulatory bodies to follow Japan's lead. This is a bad move. Regulatory agencies around

the world should resist pressure to create such fast-track systems, at least until Japan has proved that its system works. That will take time. The country will have to demonstrate that its health-care system can withstand the costs of the new regenerative-medicine treatments, and that patients do not feel cheated. What happens when, inevitably, one of the fast-track drugs turns out to be ineffective? Company officials and government representatives say that patients will not be reimbursed, even though some might have paid up to ¥4.5 million (the rest covered by health insurance) for an ineffective treatment.

Japan's drug authority must guarantee that the post-commercialization evaluation of the drugs will be as rigorous as it says. It will not be easy to rein in a drug that has already been approved, whether that approval is conditional or not. If lax evaluation means that ineffective drugs are not revealed, or are not taken out of circulation, Japan could find itself flooded with unsuccessful treatments. And that would not be good for patients, the government or the biotech companies that want to see their truly effective medicines noted as such. ■

**"Japan could find itself flooded with unsuccessful treatments."**

# Future-proofing

*Hard decisions on issues that will affect future generations should not be sidestepped.*

**"I**t was the best of times, it was the worst of times, it was the age of wisdom, it was the age of foolishness." Charles Dickens had it about right in *A Tale of Two Cities*. As *Nature* went to press, negotiators in Paris were edging towards a global deal to try to secure a safe ecological future for all — a few weeks after mass murder on the city's streets. Nobody was getting too excited about the prospects, or the impact of an eventual deal, but those at the meeting seemed confident that nations would come together to agree, well — something. From a political perspective, a weak treaty that nudges action against climate change forwards is wiser than nothing at all. From a scientific point of view, of course, anything less than full speed ahead is foolishness.

Meanwhile, a week ago and a world away in Washington DC, scientists were meeting to discuss another future for the world. Assuming that the climate talks can secure a habitable planet for humanity, then just what will those humans be like? While environmentalists search for new technologies to safeguard the future, biologists have a whole box of new tools that can reveal and manipulate the genome. As we report on page 173, the atmosphere at the Washington meeting — convened to discuss the implications of human-gene-editing techniques — was cordial and hopeful.

The parallels between the two issues — global warming that can alter the world outside and technology such as CRISPR–Cas9 that can rewrite our world inside — are telling. Most of the major concerns will

not affect the people currently worrying about them. They are talking and acting on behalf of generations to come, those unspoken voices that trouble us from the future. Is it fair to leave them an ecosystem very different from the one we enjoy, which they will recognize only as history? Is it ethical to fiddle with the human germ line to introduce changes that will echo through future families and alter the legacy of human diversity?

Politicians and policymakers struggle when they are required to put the needs of the unborn ahead of the demands of voters and funders. So both the climate negotiations and the gene-editing discussions have a zoom function, to illustrate the near-term challenges and opportunities: the local pledges and actions to cut emissions right now, and the basic research needed to make an experimental technique safe for clinical use. Both are necessary steps, but both in their own way dodge the big questions. What does the world do to accelerate these feel-good emissions cuts and gear them up to meaningful collective action? And what does society want to do with a fully operational gene-editing system?

The current discussions on genetics and climate have much to commend them. They have learned the lessons of the past and are trying to break down the conventional political and scientific hierarchy to reflect the rise of nations such as China. The people most directly affected by the decisions reached — indigenous and poor communities in the developing world and individuals and families affected by genetic disorders — are being consulted and listened to (although not enough). The mood is, generally, cordial and constructive.

The worry is that the bar in both discussions is set too low. We should be wary about celebrating times that seem the best only because we have put the worst decisions off for another day. ■

**➤ NATURE.COM**  
To comment online,  
click on Editorials at:  
[go.nature.com/xhunq](http://go.nature.com/xhunq)





## Business needs to embrace sustainability targets

*Companies that do not align their management strategies with sustainable development goals will lose out, warns Klaus Leisinger.*

In September, countries adopted the United Nations' 2030 Agenda for Sustainable Development: goals and targets to end poverty, protect the planet from degradation, ensure prosperity and foster peaceful, just and inclusive societies. This was historic — never before have world leaders agreed on such a broad agenda to transform societies.

Of course, the real success of Agenda 2030 demands implementation by all of us, from governments and other societal actors to individuals. Companies — especially large, international corporations that set the standards for industry sectors and markets — have a pivotal role and must go further than 'business as usual'. Society's rules have changed and managerial attitudes must keep up. Here is how.

First, companies must buy into the Agenda 2030 values: freedom, equality, solidarity, tolerance, respect for nature and shared responsibility. Politically convenient lip service crafted by their communications specialists is insufficient. The values must be built into strategies, business models and product development. Corporations that ignore ecological, social and economic knowledge and the global values expressed by the international community show a lack of integrity.

New, integral thinking is crucial. If a corporate culture focuses only on short-term economic key performance indicators, then so will employees, and they will choose the cheapest suppliers irrespective of whether they have degrading working conditions or destructive environmental practices. The Volkswagen emissions scandal demonstrates what can happen if management remains distant from its employees.

Corporate transformation for sustainability is a leadership task. Just as stairs are cleaned from the top down, so institutional reforms must be lived by executives, not just proclaimed. I have spent the past three decades at the helm of a major corporate foundation and engaging with civil society, governments and UN institutions. I know of no companies that are doing enough today to meet the sustainable goals. Their intergenerational and integrated nature requires difficult, radical shifts. 'Doing no harm' as suggested by current corporate responsibility efforts — fair labour norms, the precautionary principle and UN Guiding Principles on Business and Human Rights — is no longer enough.

Most of the business leaders I know are good people and managers. Few are moral heroes. The new UN goals, they tell me, are idealistic, and they agree with them in principle. But there are many reasons for them not to act, at least not immediately — too much existing regulation; foreign competitors with cost advantages; shareholders; and financial analysts who will not accept lower profits even for a short time. What is in it for these managers? They know that customers, even those

who say that they agree with the sustainable vision, will not pay the prices that tell the ecological and social truth.

Truly sustainable companies must embrace emissions reductions and green energy, and control their use of non-renewable resources. They must invest in business models that bring essential products and services within the reach of low-income families, stop selling goods that are not biologically degradable, and develop alternatives that are. And they must have the same environmental and social standards at all production sites.

All these reforms have immediate costs and their returns will not flow quickly. Some sustainable pioneers will identify opportunities and enjoy first-adopter advantages — car companies that invested in hybrid solutions are favoured by ecologically conscious customers, for example — but, yes, corporate sustainability investments will increase costs,

and might reduce sales and lower profits, at least in the short term.

This is why we must look at such reforms as the necessary price of integrity, an investment in credibility and an insurance premium for future success. In time, costs will be offset by success in new markets, greater motivation of employees and customers, and reputational gains. Values such as ecological sustainability, social inclusion and human rights must be assessed alongside measures of economic success when decisions are made. Codes of conduct and corporate sustainability guidelines must be amended. Performance appraisals and bonus decisions must refer to these criteria.

Sustainability thinking must penetrate the entire corporate value chain. That starts with the sourcing of raw materials and services, through to transport, employment practices and environmental stewardship in production. It extends to packaging and delivering, the use of products and services by customers, and product disposal, reuse or recycling. All new investments and research and development should undergo a sustainability assessment similar to the one proposed by the UN Guiding Principles on Business and Human Rights.

Communications also need an overhaul. An enlightened public does not find corporate responsibility 'poetry' — pictures of smiling children and jolly workers — credible. Communications must be truthful and show transition as a learning process, with setbacks and obstacles reported as well as successes and progress.

Competing with integrity is smart risk management. Acting against global societal interests results in reputational damage, law cases, penalties and more regulation — as banks around the world and Volkswagen are finding out. ■

**Klaus Leisinger** is president of the Foundation Global Values Alliance and professor of sociology at the University of Basel, Switzerland.  
e-mail: [klaus.leisinger@globalvaluesalliance.ch](mailto:klaus.leisinger@globalvaluesalliance.ch)

VALUES  
MUST BE  
BUILT INTO  
**STRATEGIES,**  
BUSINESS  
MODELS AND  
**PRODUCT**  
DEVELOPMENT.

➔ **NATURE.COM**  
Discuss this article  
online at:  
[go.nature.com/xs5fja](http://go.nature.com/xs5fja)



# RESEARCH HIGHLIGHTS

Selections from the  
scientific literature

## GENOME EDITING

### Editing enzyme made more precise

By tweaking the structure of an enzyme that cuts DNA, researchers have lowered the error rate of CRISPR–Cas9 gene editing.

Feng Zhang at the Broad Institute of MIT and Harvard in Cambridge, Massachusetts, and his colleagues engineered the Cas9 enzyme so that it is less likely to act at DNA sites that are not targeted by the RNA molecules that guide the enzyme. The team generated several versions of Cas9 that reduced off-target errors by at least tenfold compared with unaltered Cas9 enzymes. Three of those versions were just as active at their target sites as ordinary Cas9.

Even so, the error rate may need to be reduced further for CRISPR gene editing to be used for human therapies.

**Science** <http://doi.org/9q2> (2015)

## MICROBIOMES

### Gut bacteria change with cold

Gut microbes alter the metabolism of mice to help the animals to adapt to the cold.

Mirko Trajkovski of the University of Geneva, Switzerland, and his team studied mice that were kept at low temperatures, and found that their microbiomes diverged from those of mice housed at room temperature. When microbes from cold-exposed mice were transplanted to those without any bacteria, the recipients made more beige fat cells, which burn more energy to produce heat than white fat cells do. The transplants also improved the recipients' metabolism by boosting their insulin sensitivity. Both cold

exposure and transplants increased the absorptive surface of the gut, enhancing caloric uptake.

The microbiome's control over energy expenditure could make it a therapeutic target for combating obesity, the researchers suggest.

**Cell** 163, 1360–1374 (2015)

## GEOLOGY

### Mediterranean quake risk rises

Earthquakes threaten the eastern Mediterranean more than scientists had thought.

A team led by Vasiliki Mouslopoulou at the GFZ

German Research Centre for Geosciences in Potsdam analysed and carbon-dated ancient shorelines in Crete to estimate when and how quickly earthquakes over the past 50,000 years pushed the island upwards. They conclude that at least 40 quakes of magnitude 8 or greater would have raised the land to its current level.

The quakes came in clusters separated by quieter periods. The entire eastern Mediterranean could be at higher risk of future mega-quakes — and the resulting tsunamis — than previously suspected.

**Geophys. Res. Lett.** <http://doi.org/9q3> (2015)



## ZOOLOGY

### Cuttlefish use electric camouflage

When cuttlefish freeze as enemies approach, they are reducing their bioelectric fields, which predators can detect.

Christine Bedore at Duke University in Durham, North Carolina, and her colleagues showed videos of approaching predators, including sharks, to *Sepia officinalis* cuttlefish (pictured) in lab tanks, and analysed their behaviour and the electric fields they emitted. Most cuttlefish froze when they saw the predator, flattening themselves against the tank, reducing

their breathing rate and closing their orifices. These behaviours reduced their bioelectric fields, which are generated by ion exchange between respiratory structures and seawater.

Captive sharks that were presented with electric currents mimicking both frozen and resting cuttlefish were more likely to bite at the stronger fields that represented resting creatures. This suggests that the freeze response lowers the electric 'visibility' of cuttlefish.

**Proc. R. Soc. B** 282, 20151886 (2015)

## SYNTHETIC BIOLOGY

### Kill switches limit modified microbes

Genetic switches that make bacteria self-destruct in response to specific signals could help to keep engineered microbes in check.

Microbes that have been genetically modified for industrial and other uses rely on the addition of certain nutrients to grow, but this makes it difficult to design organisms that can be contained in different environments. To overcome this, James Collins of the Massachusetts Institute of

P. SUTTER/CORBIS

Technology in Cambridge and his colleagues designed gene circuits that act as 'kill switches' when they are introduced into the genome of the bacterium *Escherichia coli*. The switches block essential genes or induce lethal toxins when triggered by environmental chemicals.

Such kill switches can be customized to respond to different cues, the authors say. *Nature Chem. Biol.* <http://dx.doi.org/10.1038/nchembio.1979> (2015)

## PLANT BIOTECHNOLOGY

## CRISPR clips crop genes

The CRISPR–Cas9 genome-editing system has been used to alter the genomes of two crop plants: barley (*Hordeum vulgare*; pictured) and *Brassica oleracea*, a species that includes broccoli and cabbage.

CRISPR–Cas9 allows researchers to easily engineer mutations in genomes and has been tested in some crops, including rice and wheat. Cristobal Uauy and Wendy Harwood at the John Innes Centre in Colney, UK, used the system in barley and the brassica species to knock out the function of genes encoding certain plant hormones that are involved in growth and seed development — both important agronomic traits.

The team generated heritable mutations and the modified plants contained no foreign genes. However, the editing system occasionally introduced unwanted, off-target genetic changes.

*Genome Biol.* 16, 258 (2015)



HAGEN ZAPF/IMAGEBROKER/CORBIS

## PHYSICS

## Long delay for electron decay

An underground experiment has yielded the strongest evidence so far that electrons are stable, by showing that they last for at least 66 billion billion years before decaying into photons and neutrinos.

The Borexino detector, located underground at the Gran Sasso National Laboratory in central Italy, is looking mainly for neutrinos. The electron-stability result it produced is about 100 times longer than previous limits, and was obtained thanks to the extremely low natural radioactivity of the 300 tonnes of highly purified mineral oil at its core, according to Mauro Pallavicini of the University of Genoa, Italy, and his team.

Because the decay of negatively charged electrons into neutral particles seems even less likely than before, the results also confirm the fundamental law saying that net electric charge is conserved. In principle, however, electrons could still decay into some other unknown particles, the authors say.

*Phys. Rev. Lett.* 115, 231802 (2015)

## LASER PHYSICS

## Lasers trigger X-rays efficiently

Bombarding clouds of a noble gas with an intense UV laser emits X-ray pulses that could be useful for imaging applications.

Single atoms zapped by a UV laser emit bright X-rays, but to get such X-rays from a cloud of atoms, a longer-wavelength laser is needed. To maximize the intensity of the emitted X-rays, Tenio Popmintchev at the University of Colorado Boulder and his colleagues found UV wavelengths that elicit strong emissions over a wide range of frequencies from single atoms or ions and clouds

## SOCIAL SELECTION

Popular topics on social media

### Should DNA donors see genomic data?

Geneticist Charles Danko turned to Twitter this week to ask for help in convincing his institutional review board (IRB) at Cornell University in Ithaca, New York, that he should be allowed to let his study participants download their own genetic information. He tweeted: "Writing IRB amendment to give subjects their genomic data. IRB does not like. Suggested reading to make the case?" Several scientists responded to Danko's request, including computational geneticist Yaniv Erlich at Columbia University in New York City, who pointed out that direct-to-consumer (DTC) genetics companies such as 23andMe already give people their data. "Another argument is that ~2 million people have access to their raw genetic data through DTC," he wrote. "No adverse events have been reported so far due to access to raw data. Other research projects do the same."

➔ **NATURE.COM**  
For more on popular papers:  
[go.nature.com/bxjd8q](http://go.nature.com/bxjd8q)



of these particles. By carefully choosing a laser's colour and shaping a guide for the light, the team created intense X-ray pulses just 100 attoseconds long (1 attosecond is  $1 \times 10^{-18}$  seconds).

The technique produces light that could be powerful enough to probe the details of electron motion, the authors report. *Science* 350, 1225–1231 (2015)

## PHOTONICS

## Seeing movement around corners

A camera can track the movement of an object hidden around a corner.

Researchers can already reveal images of hidden items by firing a laser onto the floor and studying the light that is scattered back from all surfaces, including concealed ones. But such systems

work too slowly to track any movement. Genevieve Garipey and Daniele Faccio at Heriot-Watt University in Edinburgh, UK, and their colleagues used a laser and an ultrafast camera to measure the position and arrival time of photons from the laser bouncing back off a figurine placed around a corner (pictured). This allowed them to track the object, even while it was moving at a speed of a few centimetres per second.

The method could be useful for remote observation in hazardous areas, or to prevent vehicles colliding at blind corners, the authors say. *Nature Photon.* <http://dx.doi.org/10.1038/nphoton.2015.234> (2015)

➔ **NATURE.COM**  
For the latest research published by Nature visit:  
[www.nature.com/latestresearch](http://www.nature.com/latestresearch)



# SEVEN DAYS

The news in brief

## POLICY

### Anti-ageing trial

The US Food and Drug Administration has informally agreed to allow researchers to test whether the diabetes drug metformin prolongs healthy lifespan in humans, said leaders of a proposed clinical trial on 30 November. Physician Nir Barzilai of the Albert Einstein College of Medicine in New York City and his colleagues plan to study the drug in individuals who have, or are at risk, of age-related disease. Details of the trial, called Targeting Aging with Metformin, are still being formalized.

### Gene-editing talks

Gene-editing technologies should not be used to modify human embryos that are intended for use in establishing a pregnancy, declared a 3 December statement prepared by a global conference on genome editing. The International Summit on Human Gene Editing, a three-day meeting at the US National Academy of Sciences in Washington DC, also called for cautious development of gene-editing applications that cannot be passed on to offspring — such as correction of the mutations that cause sickle-cell disease or modification of immune cells to target cancer. See page 173 for more.

## EVENTS

### Draft climate deal

On 5 December, negotiators at the United Nations climate talks in Paris released a draft agreement to reduce global greenhouse-gas emissions. The 48-page document will be debated by government ministers, with the aim of agreeing a new global climate deal by the scheduled end of the talks, on 11 December.



GALAPAGOS NATL PARK

## Lonesome George was not alone

The Galapagos National Park (GNP) is attempting to restore two lineages of extinct giant tortoise, including the Pinta tortoise species that went extinct when Lonesome George died in 2012. Many of the tortoises on Wolf Volcano, on Isabela Island, are known to be hybrids with mixed ancestry, and on 4 December the GNP announced that its

rangers had brought 32 of these tortoises into captivity (pictured). Two of these, a male and a female, are not close relatives of George, but they share some genes with his Pinta species. Five animals are hybrid descendants of the Floreana tortoise, which vanished in the nineteenth century. See [go.nature.com/tqsidx](http://go.nature.com/tqsidx) for more.

One of the big questions is how countries' progress on meeting their pledges will be reported and reviewed. The United States has pushed for better monitoring of emissions by developing countries, but many of these nations worry about the cost and feasibility of doing this. See [go.nature.com/coante](http://go.nature.com/coante) for more.

### Ivory price falls

The price of ivory in China has plummeted, raising hopes for a reduction in elephant poaching. Raw ivory fetched an average of US\$2,100 per kilogram in Beijing in 2014, but the price last month was \$1,100 per kilogram, according to independent researchers Lucy Vigne and Esmond Martin,

working with charity Save the Elephants. China has attempted to clamp down on its domestic ivory trade, and celebrities have highlighted the damage to elephant populations. A drop in price could lead to fewer elephants killed in Africa, says Save the Elephants.

### Red smog alert

On 7 December, Beijing issued its first red alert for air pollution under a system put in place in 2013. The red alert, for heavy pollution expected to last for three or more days, was planned to be in effect until midnight on 10 December. Some industrial plants and construction sites shut down and schools were advised to close. When the

alert came into effect, one measure of air quality, the level of particulate matter less than 2.5 micrometres in diameter, was nearly 300 micrograms per cubic metre. The World Health Organization says that levels greater than 25 micrograms can start to affect health.

## RESEARCH

### Australian rankings

The quality of research in Australia has improved, according to the results of a nationwide audit published on 4 December. The 2015 Excellence in Research for Australia exercise found that 35% of university departments submitted for assessment had improved

on the results obtained in a similar 2012 exercise, with 32% of submissions being judged “well above world standard” compared with 18% three years ago. An unofficial ranking by *The Australian* based on the audit judged the Australian National University in Canberra to be the country’s top institution.

## FACILITIES

## Telescope blow

Hawaii’s supreme court dealt a major blow to the Thirty Meter Telescope (TMT) project on 2 December, when it revoked the permit allowing the mega-telescope to be built near the summit of Mauna Kea. Many Native Hawaiians have protested against construction on what they regard as a sacred mountain. The court said that Hawaii’s Board of Land and Natural Resources should not have approved the permit in 2011, because it did so before protestors could air their side in a contested case hearing. TMT officials must now apply for a new construction permit if they wish to proceed. See page 176 for more.

## Base on skis

The British Antarctic Survey has started planning the process of moving its Halley VI research station (pictured), which is



threatened by a crack in the ice 7 kilometres away that could cut it off from the rest of the shelf. The station consists of 8 connected pods resting on skis on the 150-metre-thick Brunt Ice Shelf. It is designed to be transportable, but the move will be the first since the station became operational in 2012, and comes earlier than expected. In July 2014, during the Antarctic winter, the base temporarily lost power as temperatures outside plummeted to  $-55^{\circ}\text{C}$ . The move is expected to be completed in the 2016–17 Antarctic research season.

## FUNDING

## Facebook fund

Health research is among the priorities of a new foundation to be created by Facebook founder Mark Zuckerberg and

his wife, Priscilla Chan. In an open letter on 1 December to their daughter, Maxima, the couple announced that they will invest most of their wealth in a limited-liability corporation called the Chan Zuckerberg Initiative, dedicated to improving health and education. The letter names as priorities disease prevention, heart disease, cancer, stroke and neurodegenerative and infectious diseases.

## Australian cash

The Australian government on 7 December announced a 4-year, Aus\$1.1-billion (US\$790-million) National Innovation and Science Agenda. The pot includes Aus\$106 million of tax incentives for early-stage ‘angel’ investors, Aus\$75 million for

## COMING UP

## 14–18 DECEMBER

Scientists gather in San Francisco, California, for an American Geophysical Union meeting — and an advance screening of the latest Star Wars film. [fallmeeting.agu.org/2015](http://fallmeeting.agu.org/2015)

## 15–20 DECEMBER

Chemists head to Honolulu, Hawaii, for the Pacificchem congress. [www.pacificchem.org/](http://www.pacificchem.org/)

a digital and data productivity network, and Aus\$127 million for research grants. The government also made a Aus\$1.5-billion promise to secure funding for the country’s threatened research infrastructure over the next 10 years. National research agency the CSIRO, which last year suffered a Aus\$111-million funding cut, will get Aus\$70 million towards a Aus\$200-million innovation fund.

## PEOPLE

## ORI head named

Neuroscientist Kathryn Partin has been made head of the US Office of Research Integrity (ORI), according to an article in the *Report on Research Compliance* on 3 December. The ORI oversees investigations into misconduct concerning researchers funded by the US Department of Health and Human Services. Partin is currently assistant vice-president for research at Colorado State University in Fort Collins, where she heads the university’s research-integrity office. She replaces interim director Don Wright, who has led the agency since previous director David Wright resigned in March 2014.

► [NATURE.COM](http://NATURE.COM)

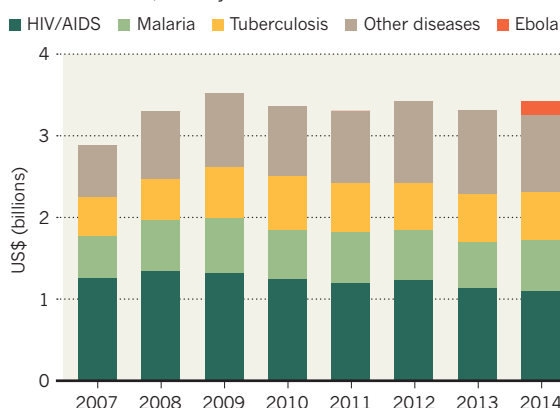
For daily news updates see: [www.nature.com/news](http://www.nature.com/news)

## TREND WATCH

Money spent on neglected-disease research rose by US\$150 million in 2014 to \$3,377 million, says the latest G-FINDER report from analysts Policy Cures in Sydney, Australia. Investments made to tackle the 2014 Ebola epidemic in West Africa were responsible for the rise. Apart from Ebola, three diseases received the most funding: HIV/AIDS, malaria and tuberculosis, as in previous years. But removing Ebola investments from the figures shows that public funding for neglected diseases is at a seven-year low, says the report.

## FUNDING BOOST FOR NEGLECTED DISEASES

Funding in research and development for neglected diseases increased in 2014, but only because of investments for Ebola.





# NEWS IN FOCUS

**BIOTECHNOLOGY** Views on human gene editing divide world **p.173**



**SPACE** Japan's Venus probe enters orbit — five years late **p.174**

**OCEANOGRAPHY** Imploded deep-sea robot replaced by lower-risk tech **p.176**

**BIOMEDICINE** The controversy over use of aborted fetal tissue in research **p.178**

ABDELHAK SENNA/AFP/GETTY



Restrictions on harvests and exports of *Gelidium* seaweed in Morocco have affected the global supply of the lab reagent agar.

## MICROBIOLOGY

# Lab staple agar runs low

*Dwindling seaweed harvest imperils reagent essential for culturing microbes.*

BY EWEN CALLAWAY

Microbiology's most important reagent is in short supply, with potential consequences for research, public health and clinical labs around the world.

Agar — the seaweed-derived, gelatinous substance that biologists use to culture microbes — is experiencing a global downturn, marine biologists, agar producers and industry analysts told *Nature*. “There’s not enough seaweed for everyone, so basically we are now reducing our production,” says

Pedro Sanchez, deputy managing director of Industrias Roko in Polígono de Silvota, Spain, which processes seaweed to make some 40% of the world’s agar.

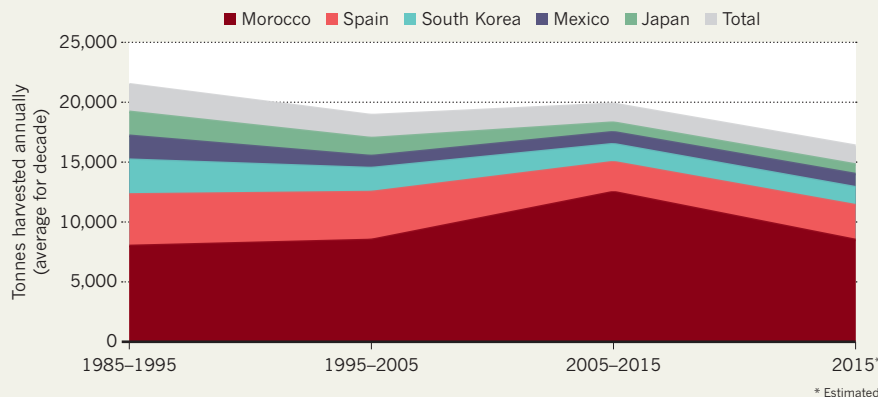
The shortage can be traced to newly enforced trade restrictions on the seaweed, arising from environmental concerns that the algae are being overharvested. It is unclear how deeply the dearth will hit researchers, but it has already pushed wholesale prices of agar to an all-time high of around US\$35–45 per kilogram — nearly triple the price before scarcities began. Individual researchers, who buy packaged agar

from lab-supply companies, can pay many times this amount.

One major supplier, Thermo Fisher Scientific of Waltham, Massachusetts, says that it has stopped selling two ‘raw’ agar products — agar that has not been mixed with other ingredients — until 2016, so that it can prioritize more-popular products that contain a mixture of agar and growth nutrients. The company reports that about 200 of its customers have been affected. Another major lab-supply company, Millipore Sigma in Billerica, Massachusetts, has also halted sales of raw agar, and it says that it will ►

## SEAWEED SHORTAGE

Harvests of *Gelidium* seaweed, from which the agar used in labs is made, are shrinking — particularly in Morocco, which is the world's major supplier.



► re-evaluate its supplies early next year.

Millipore Sigma blames the shortage on competition from food companies for purified agar. The global demand from food-makers, at several thousand tonnes annually, dwarfs the 900 tonnes that go to lab-supply companies.

## PRIZED MATERIAL

Formed of long chains of sugar molecules, agar is prized by microbiologists for its ability to form hard gels when mixed with water and growth nutrients. When a solution of bacteria is spread onto an agar-lined plate, individual cells grow into distinct colonies, allowing researchers to isolate each different strain in the mixture.

Adam Roberts, a microbiologist at University College London, drew attention to the agar-supply problem last month, when he tweeted a letter from Thermo Fisher announcing its suspension of some agar sales. His lab depends on the product to identify new antimicrobial compounds in soil bacteria.

Roberts managed to source raw agar from another supplier, but he says that his lab may have to begin rationing it and prioritizing some experiments over others. "It would be a bloody nightmare," he says. When a colleague at another institution heard about Roberts's agar-supply problem, she started to hoard her own stash. "If it gets more serious and hard to come by, I don't know what we'll do," says Roberts.

Since the introduction of agar plates in the 1880s — which enabled researchers to isolate tuberculosis, cholera and other

disease-causing bacteria for the first time — bacteriological agar has been derived from a clutch of red seaweed species belonging to the genus *Gelidium*.

Algae of this type grow atop rocky sea beds, forming vast underwater lawns of bushy, red fronds, and they favour cool, turbulent waters that provide a steady supply of oxygen and other nutrients — a preference that makes industrial-scale farming impossible. "It's not cultivated, and it's not possible to cultivate — although we've wasted a lot of money trying to do it in the past," says Sanchez. In some places, *Gelidium* is harvested by underwater divers or when the tides roll in, but the seaweed is most commonly collected when storms wash it ashore.

The geographical sources of *Gelidium* have shifted over the decades. Before the Second World War, Japan was king; Portugal was also once a leading supplier. Now, most of the world's agar derives from *Gelidium* grown in Morocco, with Spain in second place and Portugal, France, Mexico, Chile, South Africa, Japan and South Korea all contributing smaller quantities (see 'Seaweed shortage').

Current agar deficiencies are due mostly to an unsteady supply of Moroccan *Gelidium*, say industry insiders. Throughout the 2000s, the nation regularly harvested as much as

14,000 tonnes per year, which was sold on to foreign and domestic agar producers. But citing concerns over dwindling *Gelidium* populations, the Moroccan government cut the legal annual harvest to around 6,000 tonnes, and has limited foreign exports of the algae to around 1,200 tonnes. Although the changes were imposed in 2010, the country only began to enforce these trade limits last year, says Sanchez.

## GOLD RUSH

There is evidence, says marine ecologist Ricardo Melo at the University of Lisbon, that Moroccan *Gelidium* stocks were being over-harvested by throngs of beach-combers in search of 'red gold,' as the seaweed is known. But Melo says that the trade restrictions make little sense from a conservation point of view.

The domestic Moroccan market is now flooded with *Gelidium*, while the rest of the world struggles with a massive shortage. This has benefited Morocco's lone agar producer, which can now buy the seaweed at rock-bottom prices, but the move has vastly increased the cost for producers elsewhere.

And that means that companies such as Thermo Fisher and Millipore Sigma, which buy purified agar from producers and sell it on to researchers as packaged products, have little choice but to pay the skyrocketing prices that agar now commands, says Dennis Seisun, who runs an industry-analysis firm called IMR International in San Diego, California. "The ones that had contracts with the suppliers are getting preferential treatment, but I'm pretty sure they're not getting all they want."

The European Commission has complained to the Moroccan government that the country's export restrictions violate its free-trade agreements with European Union countries, but to little avail. And prospects for new sources of *Gelidium* are bleak. Unexploited populations of the seaweed exist off the shores of North Korea, but "it's not easy to work in a country like North Korea," says Sanchez.

Agar substitutes, such as a seaweed product called carrageenan, have proved unsuitable for culturing microbes. Before a technician introduced agar to his lab, pioneering German microbiologist Robert Koch isolated bacteria on potato slices. "Unless we go back to what Koch used to do — use a potato," Roberts says, "there's no real alternative." ■

PEDRO SANCHEZ/INDUSTRIAS ROKO



## TOP NEWS



Daily news, commentary and video updates from the Paris climate talks [go.nature.com/cslgig](http://go.nature.com/cslgig)

## MORE NEWS

- Pirate research-paper sites play hide-and-seek with publishers [go.nature.com/7ilnal](http://go.nature.com/7ilnal)
- Global greenhouse-gas emissions set to fall in 2015 [go.nature.com/ekk77m](http://go.nature.com/ekk77m)
- Mosquitoes engineered to pass down infertility genes [go.nature.com/vppzwb](http://go.nature.com/vppzwb)

## NATURE PODCAST



Ceres under the microscope, fetal tissue in science and the wasting disease that occurs with cancer [nature.com/nature/podcast](http://nature.com/nature/podcast)



## BIOTECHNOLOGY

# Global summit reveals divergent views on human gene editing

*Representatives discuss the ethical, social and legal issues that unite and divide them.*

BY SARA REARDON

When nearly 500 scientists, ethicists, legal experts and advocacy groups from more than 20 countries came together in Washington DC last week to produce guidelines for the use of gene editing in humans, the meeting served as a potent reminder of how far genetic engineering has permeated society.

In 1975, a group of mostly US scientists met at an iconic conference in Asilomar, California, and set stringent guidelines for moving forward with powerful new research tools that enabled the mixing of DNA between species. Forty years later, it took a much more diverse group to reach a much less definitive agreement: a recommendation not to stop human-gene-editing research outright, but to refrain from research and applications that use modified human embryos to establish a pregnancy.

Held on 1–3 December, the International Summit on Human Gene Editing was organized by the US national academies of sciences and medicine, the Royal Society in London and the Chinese Academy of Sciences (CAS). The meeting highlighted China's emerging prominence in genomics; much of the discussion surrounded an April publication by Chinese researchers who used the gene-editing technology CRISPR–Cas9 to modify a gene in non-viable human embryos (P. Liang *et al. Protein Cell* 6, 363–372; 2015).

A position statement released at the end of the meeting by its organizers did not condemn such experiments. But it said that a host of ethical and safety issues should be resolved before embryos are modified for clinical applications.

Many countries already limit research using human embryos. China explicitly bans the implantation of genetically modified embryos in a woman. Some countries ban gene editing of human embryos even in the lab. *Nature's* editor-in-chief, Philip Campbell, said in a presentation that *Nature* journals have rejected papers on human germline editing because of “non-compliance with local regulations”.

Cultural differences are likely to perpetuate the diversity of regulations. Bioethicist Renzong Qiu of the Chinese Academy of Social Sciences noted that in the United States, a politically charged debate on whether embryos have human rights has resulted in laws that ban the use of public funds for research in which



YORGOS NIKAS/SPL

The genetic modification of human embryos (orange cells) is a controversial topic.

embryos are created or destroyed. This aspect is not even part of the discussion in China, he said: “According to Confucius, human being is only after birth.”

Ephrat Levy-Lahad, a cancer researcher at the Hebrew University of Jerusalem, said that Israel is likely to welcome clinical use of genetically modified embryos. The government, which encourages large families, already pays for parents using *in vitro* fertilization to screen their embryos for genetic mutations — a technique known as preimplantation genetic diagnosis, which is simpler than genome editing and can prevent many, but not all, inherited disorders.

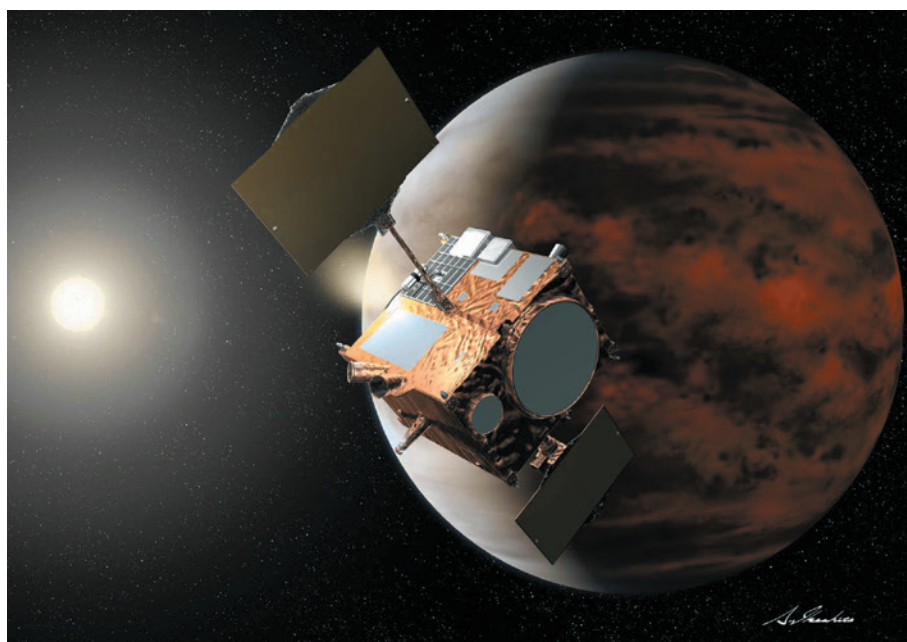
The organizers of the Washington meeting reached not just across national borders, but also across disciplinary boundaries. Social scientists and ethicists raised concerns that altering human genomes could create inequality and discrimination: in the distant future, wealthy parents could choose their child's skin colour, for instance. And sociologist Ruha Benjamin at Princeton University in New Jersey worries that the technology will create friction over which traits should be considered disorders — many deaf people, for example, do not think of themselves as disabled and want their children to share in deaf culture.

Jinghua Cao, deputy director-general of the CAS's Bureau of International Cooperation, said that such preferences were new to him.

“That is too far away from a normal way of looking at things in China,” he said. “It is good for us to learn that perspective; nevertheless, that goes too far.”

Despite differences about how far to go in applying gene editing to the unborn, nearly everyone at the meeting agreed that efforts to use gene editing after birth to correct defects in non-reproductive cells should continue. Qi Zhou, a developmental biologist at the CAS's Institute of Zoology in Beijing, said that he was surprised to learn how quickly Western countries have moved the technology into the clinic. Trials of gene editing to treat leukaemia, HIV and haemophilia are planned or under way in the United States and the United Kingdom. No specific regulatory pathway for such treatments exists in China, Zhou said, and he plans to begin discussing this with the government as soon as possible.

For all the diversity at the meeting, the organizers acknowledge that it is only the first step. Many countries were not represented, and the agenda included few presentations by people with genetic disorders. When the mother of a child born with an inherited mutation took the microphone to describe tearfully how a genetic disorder had wracked her son's body with seizures throughout his six-day life, her story generated a current of emotion not often encountered at a scientific gathering. “If you have the skills and the knowledge to fix these diseases,” she said, “then frickin’ do it.” ■ SEE EDITORIAL P.164



The Akatsuki probe spent five years orbiting the Sun.

## ASTRONOMY

# Japan's Venus probe enters orbit

*Five years after a failed insertion into the planet's orbit, Akatsuki finally reaches its target.*

BY ALEXANDRA WITZE

Japan's Akatsuki spacecraft has entered orbit around Venus, five years after its first attempt failed. On 7 December, at 8:51 a.m. Japan time, Akatsuki ignited four small thruster engines for roughly 20 minutes. The tiny push was enough to nudge the probe into the pull of Venus's gravity.

As *Nature* went to press, exactly what that orbit looks like remained unclear. But mission scientists are confident that the spacecraft has at least partly redeemed itself, after a 2010 attempt to reach Venus left Akatsuki spiralling around the Sun.

"It's in orbit!" said Sanjay Limaye, a planetary researcher at the University of Wisconsin–Madison, and a participating scientist on the mission. "Everyone is very happy."

The Japan Aerospace Exploration Agency (JAXA) planned to announce the exact details of the orbit at 6 p.m. Japan time (9 a.m. London time) on 9 December. Even the best-case scenario would see Akatsuki travel a much more stretched-out orbit around Venus than originally planned. The spacecraft could range

about 500,000 kilometres from the planet at its farthest point, taking perhaps 14 or 15 days to make each orbit. Eventually, mission controllers plan to fire the thrusters again to shrink the orbit further — to about 330,000 kilometres at its farthest point. That would see it completing a circuit around Venus about every 8 days.

"It's been quite a long period of waiting," says Masato Nakamura, JAXA project manager at the Institute of Space and Astronautical Science in Sagami-hara.

Akatsuki was launched in May 2010 on a mission to study Venus's ever-changing atmosphere, which rotates at up to 100 metres per second — much faster than the planetary surface below it. The spacecraft carries five cameras, ranging from infrared to ultraviolet wavelengths, to study different atmospheric features, including the lightning that is thought to flash through Venus's acidic clouds.

All seemed well until 7 December 2010, when the spacecraft fired its main engine

to enter Venus's orbit. Unknown to mission controllers, salt had built up on a valve between a helium tank and a fuel tank, and the blockage caused a ceramic nozzle in the propulsion system to break. Akatsuki went sailing towards the Sun, rather than into orbit around Venus.

JAXA engineers spent years studying whether they could recover the mission (M. Nakamura *et al. Acta Astronaut.* **93**, 384–389; 2014). With the main engine dead, the oxidizer fuel was also useless, so mission controllers dumped 65 kilograms of fuel into space in October 2011. This made the spacecraft lighter and easier to manoeuvre, which enabled it to reach orbit with less thrusting.

The crucial engine burn involved four of the spacecraft's eight thrusters. These smaller engines are normally used to make minor adjustments to the probe's orientation, rather than major changes to its trajectory. Because the thrusters are lower power than the main engine, they needed to burn for longer than usual.

## TWILIGHT ZONE

Despite the rescue's apparent success, the spacecraft's unexpected detour might still cause problems. Because it has spent more time closer to the Sun than originally designed, Akatsuki is warmer than expected, which may have harmed some of its equipment; this could limit operations at Venus.

During its five years in deep-space wilderness, Akatsuki conducted a little science, such as transmitting radio signals to Earth through the solar corona to measure how the Sun's turbulence scatters radio waves (T. Imamura *et al. Astrophys. J.* **788**, 117; 2014). "The past five years have been a tough period for us," says team member Takeshi Imamura.

Akatsuki is scientists' only chance at seeing the planet up close for the foreseeable future. The European Space Agency's Venus Express spacecraft stopped working a year ago, after eight years of circling the planet. "The new observations from Akatsuki will both extend and complement the data we have from Venus Express, so the scientific result of the two together will be more than the sum of the two individual missions," says Håkan Svedhem, project scientist for Venus Express. NASA has put two Venus probes on its shortlist of five candidates for the next Discovery-class mission, which would launch no earlier than 2020.

JAXA has a history of nail-biting second chances. Its Hayabusa spacecraft survived a number of near-fatal incidents on the way to and from the asteroid Itokawa. But in 2003, after an extended effort to make the mission work, JAXA lost its Mars-bound Nozomi spacecraft, first to a problem with a fuel valve, and then to a solar flare that fried its electronics.

Akatsuki's move is only the second such deep-space recovery. In 2000, NASA's Near Earth Asteroid Rendezvous spacecraft made it into orbit around the asteroid Eros after a first missed attempt in 1998. ■

AKIHIRO IKESHITA/JAXA



## GENETICS AND AGEING

# Clues to ageing from short-lived fish

*The turquoise killifish bids to become a model organism for ageing research.*

BY EWEN CALLAWAY

For the turquoise killifish, ‘live fast, die young’ is no cliché. The little African freshwater fish reaches sexual maturity three weeks after hatching and dies of old age a few months later. A pair of studies now hints at the genetic basis for this abbreviated existence, providing precious information for a small but growing cadre of labs that hope that the species will help to unlock the secrets of human ageing.

Turquoise killifish (*Nothobranchius furzeri*) inhabit transient ponds that fill up during the rainy season in Mozambique and Zimbabwe. During the brief wet period, the fish rapidly develop, mate and lay eggs that go into suspended animation during the dry season, to hatch when the rains return the following year.

A favourite of fish hobbyists, killifish are now found in dozens of labs. Elderly killifish — a couple of months old — show hallmarks of ageing. Their bright scales fade and their cognition wavers; many develop tumours. Lifespan-altering experiments that take years in mice and decades in primates can be over in months in killifish, which are also more closely related to humans than are fruit flies, nematodes and other short-lived lab organisms popular in ageing research.

“It turns out to be the shortest-lived vertebrate that can be raised in captivity,” says Dario Valenzano, a geneticist at the Max Planck Institute for Biology of Ageing in Cologne, Germany, and a member of one of two teams that have sequenced its genome. Papers from both teams appeared in *Cell* on 3 December<sup>1,2</sup>.

The turquoise killifish genome contains several clues to its peculiar, fleeting life. Valenzano and his colleagues found that variations in genes involved in nutrient sensing, DNA repair and ageing have been selected for during its evolutionary history. Such genes might prove instructive for ageing in longer-lived animals. One such is *IGF1R*, which has been linked to extreme longevity in bowhead whales, naked mole-rats and Brandt’s bat. Genes linked to *IGF1R* vary between an extremely short-lived killifish lab strain and a wild variety that can live for twice as long. The two strains also

differ in a gene that has been linked to dementia in humans.

“Maybe these genes are central hubs for regulating survival,” says Valenzano. “In some species they can accelerate ageing, and in some they can slow it down.”

The ability to enter suspended animation (known as diapause) may also be linked to lifespan; genes involved in a similar stasis in nematodes influence ageing too. In the other study<sup>2</sup>, a team led by Christoph Englert and Matthias Platzer at the Leibniz Institute on Ageing–Fritz Lipmann Institute in Jena, Germany, found substantial overlap in the expression patterns of genes in the killifish embryos in diapause and those in the brains of aged killifish.

Experiments that create fish lacking particular genes, for example, are needed to confirm whether the genes pinpointed in these studies truly influence ageing. These tests are already under way,

says Anne Brunet, a geneticist at Stanford University in California and an author on one of the studies. Earlier this year, Brunet’s team used CRISPR–Cas9 genome editing in ‘proof-of-principle’ experiments to alter several ageing-related genes in killifish<sup>3</sup>. “We are excited at trying to make it live longer,” she says. Her team is also screening drugs in killifish to see if any lengthen its lifespan or slow tissue decay.

Killifish hold much promise for ageing research, says Matt Kaeberlein, a biochemist at the University of Washington in Seattle, but they have not yet yielded the insights that have come out of simpler organisms. Scientists have a long list of genes that extend or shorten lifespan in nematodes, fruit flies and yeast, he notes, but it remains to be seen whether genes such as *IGF1R* have these pivotal roles in killifish. “It’s a little bit early to know for sure whether the killifish is going to emerge as the next greatest model of ageing.” ■

1. Valenzano, D. R. *et al. Cell* 163, 1539–1554 (2015).

2. Reichwald, K. *et al. Cell* 163, 1527–1538 (2015).

3. Harel, I. *et al. Cell* 160, 1013–1026 (2015).

## ASTRONOMY

# Mega-scope's permit revoked

*Court decision throws Thirty Meter Telescope into limbo.*

BY ALEXANDRA WITZE

Hawaii's supreme court has ruled that the construction permit for the Thirty Meter Telescope (TMT) on top of the mountain Mauna Kea is invalid. The decision on 2 December is a major blow to the international consortium backing the US\$1.5-billion telescope, and a win for the Native Hawaiians who have protested against its construction on what they regard as a sacred summit.

Hawaii's Board of Land and Natural Resources should not have approved the permit in 2011, the court said, because it did so before protestors could air their side in a contested case hearing. "Quite simply, the Board put the cart before the horse when it issued the permit," the court decision reads. "Accordingly, the permit cannot stand."

"TMT will follow the process set forth by the state, as we always have," TMT board chair Henry Yang said in a statement. "We are assessing our next steps on the way forward."

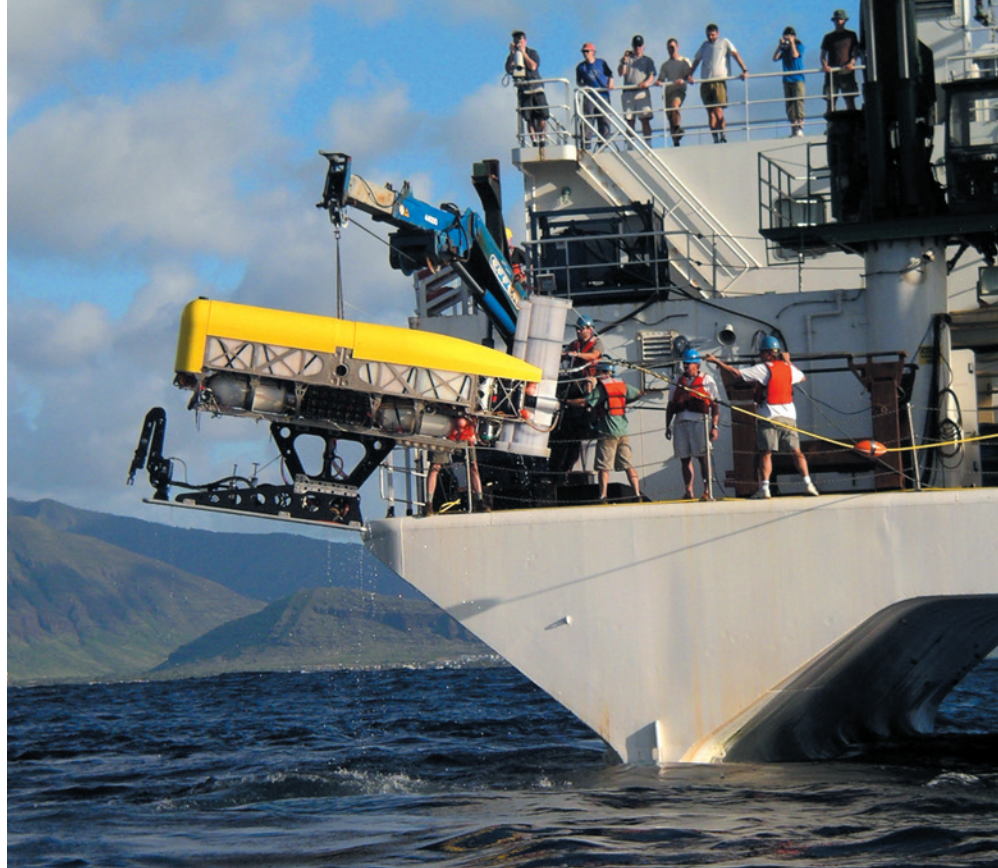
It is unclear whether and how the TMT will proceed given the new ruling. Work on the telescope's components has continued at sites outside Hawaii, but the court's decision to block the construction permit is a significant setback. To restart work on Mauna Kea, the project would have to acquire another permit from the board.

Part of the planned TMT site has been cleared, and construction was to have begun last April. But protestors have blocked the roads to the site and pursued legal means to halt the project.

Thirteen observatories — one with multiple telescopes — currently sit on Mauna Kea in a science reserve operated by the University of Hawaii. One existing telescope is being dismantled and two others are slated for decommissioning, after the fight over the TMT accelerated plans to limit development on the mountain top.

The skies above Mauna Kea are among the clearest in the world. Some Native Hawaiians say that the benefits to astronomy do not outweigh the need to respect and protect the natural and cultural environment. Many took to social media to praise the court's decision.

The TMT's partners are the University of California and the California Institute of Technology, along with research entities from the governments of Canada, China, India and Japan. ■



Nereus (yellow) was the only vehicle in the world capable of reaching the deepest parts of the ocean.

## MARINE SCIENCE

# Ocean-diving robot will not be replaced

*Woods Hole Oceanographic Institution will spend insurance money for Nereus vehicle on lower-risk projects.*

BY DANIEL CRESSEY

When the underwater robot Nereus imploded at sea more than a year ago, oceanographers were left without a vehicle that can reach the deepest parts of the ocean. Now Nereus's operator has told *Nature* that it will not replace the submersible.

The Woods Hole Oceanographic Institution (WHOI) in Massachusetts says that it will instead spread the insurance money for Nereus across multiple, lower-risk projects.

Some oceanographers say that they will miss Nereus's unique exploration capabilities, but other efforts to build similar robots that can reach the very bottom of the sea are afoot in the United States and China.

WHOI originally built Nereus at a cost of around US\$8 million — which includes its design, development and testing — with funding from the US National Science Foundation, the Office of Naval Research

and the National Oceanic and Atmospheric Administration. Among other things, the institute hoped to use the robot, which was a 'hybrid' capable of being controlled remotely and operating autonomously, to investigate the ocean's hadal zone.

This area, in deep-sea trenches between 6,000 and 11,000 metres down, is one of the least explored regions on Earth. Exactly which organisms live down there, how they survive and how they might be altered by pressures such as climate change and pollution are still only poorly understood. Any research vehicles operating at such depths must withstand intense pressure from the weight of the water above, so they are expensive to make and prone to accidents.

WHOI lost contact with Nereus during a dive in the Pacific Ocean in May 2014 — probably because a failure in one of its sealed buoyancy spheres, or in the housing around a piece of equipment, set off a catastrophic implosion at a depth of some 10,000 metres (see *Nature* 509, 408–409; 2014).



At first, researchers hoped that the institute would build a replacement vehicle. But Andy Bowen, an engineer and director of WHOI's National Deep Submergence Facility, told *Nature* that after weighing up the risks and benefits, the institute decided that the money would be better spent on less risky projects.

The \$3 million insurance payout will go towards a Nereus legacy fund to support activities “in keeping with the spirit of Nereus”, he says.

This includes developing technology to improve WHOI's undersea vehicles that do not go as deep as Nereus — as well as deep-sea ‘landers’, which go to full depth but are unable to move around. They simply sink to the bottom with various pieces of equipment on board, and are later recovered.

Such landers are the only tools currently available to explore the hadal zone — and they are no substitute for submersibles, says Jeffrey Drazen, a deep-sea researcher at the University of Hawaii at Manoa. “You put a lander down and hope for the best,” he says. By contrast, Nereus could travel around under water, relaying a real-time video feed, and be moved in response to observations. “We need both,” says Drazen.

Only having landers restricts exploration opportunities, agrees Alan Jamieson, a hadal-zone researcher at the University of Aberdeen,

UK. Last year, his team used landers to collect hundreds of hours of footage from the Mariana Trench, the ocean's deepest point, but could not take transects, a common fieldwork technique in which data is collected at multiple points along a set path. Neither could the team pick up samples, although some landers can do this.

However, Jamieson understands the choice not to recreate Nereus, describing the construction of a single, expensive robot as “a very good example of putting all your eggs in one basket”. “We need to get clever in how we access the hadal zone,” he says.

A more conservative approach has delayed — but not stopped — an effort to build a

**“We need to get clever in how we access the hadal zone.”**

full-depth vehicle at the Schmidt Ocean Institute, a private foundation in Palo Alto, California.

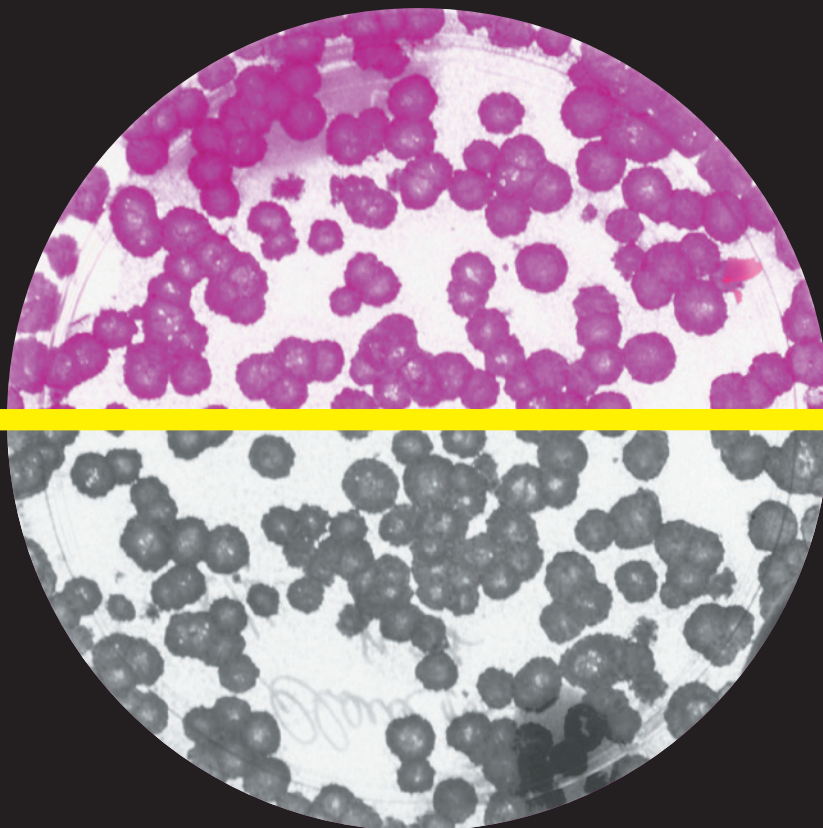
Rather than having this vehicle completely ready by 2016 as originally planned, the institute now aims to create a series of deep-sea submersibles. One will be delivered each year starting in 2016, with gradually increasing depth capabilities, says the institute's director of research, Victor Zykov. This could lead to a full-ocean-depth vehicle by 2019, if all the precursors prove successful.

“When I have discussed our plans with the deep-ocean scientists, most of them understand and appreciate this approach,” says Zykov. “They appreciate how difficult it is to operate in the ocean's deepest trenches.”

The plans in China are even bigger. In a paper published in 2014, Weicheng Cui of the Hadal Science and Technology Research Center at the Shanghai Ocean University, described a rough plan to build and deploy three landers, one robotic submersible and one human-occupied vehicle that can all operate in the hadal zone (W. Cui *et al. Meth. Oceanogr.* **10**, 178–193; 2014).

Cui told *Nature* that the first lander and the robotic submersible are currently undergoing sea trials, and that a mother ship that would control them is under construction. Around August or September 2016, he hopes to do trials in the Mariana Trench, sending the landers and the submersible to 11,000 metres. He plans eventually to use the robot to scout areas, then deploy the landers and the crewed submersible to conduct more detailed research. The project is backed by a mixture of government funding and private investment.

As to working out exactly what destroyed Nereus, says Bowen, “we'll never know — short of going and recovering the debris, which isn't financially viable. And of course there's nothing that can get there at the moment.” ■



# THE TRUTH ABOUT FETAL TISSUE RESEARCH

*The use of aborted fetal tissue has sparked controversy in the United States, but many scientists say it is essential for studies of HIV, development and more.*

BY MEREDITH WADMAN

Every month, Lishan Su receives a small test tube on ice from a company in California. In it is a piece of liver from a human fetus aborted at between 14 and 19 weeks of pregnancy.

Su and his staff at the University of North Carolina at Chapel Hill carefully grind the liver, centrifuge it and then extract and purify liver- and blood-forming stem cells. They inject the cells into the livers of newborn mice, and allow those mice to mature. The resulting animals are the only 'humanized' mice with both functioning human liver and immune cells and, for Su, they are invaluable in his work on hepatitis B and C, allowing him to probe how the viruses evade the human immune system and cause chronic liver diseases.

"Using fetal tissue is not an easy choice, but so far there is no better choice," says Su, who has tried, and failed, to make a humanized mouse with other techniques. "Many, many biomedical researchers depend on fetal tissue research to really save human lives," he says. "And I think many of them feel the same way."



An explosive climate has surrounded US research with fetal tissues since July, when an anti-abortion group called the Center for Medical Progress in Irvine, California, released covertly filmed videos in which senior physicians from the Planned Parenthood Federation of America bluntly and dispassionately discussed their harvesting of fetal organs from abortions for use in research. Planned Parenthood is a non-profit women's health provider that received US\$528 million of government money in 2014, much of it in reimbursements for services ranging from contraception to cancer screenings, which it provides largely to poor women. Abortions, which are performed at about half of Planned Parenthood's 700 clinics, constitute 3% of its services. A handful of clinics in two states supply fetal tissue for research.

The videos provoked a furore that has intensified over the past few weeks. On 3 December, the Republican-led US Senate voted to strip Planned Parenthood of government funding. This is despite the fact that fetal tissue research is legal, the US National Institutes of Health (NIH) has been funding it for decades and President Obama is sure to veto the bill, should it reach his desk. A few days earlier, on 27 November, a gunman shot dead three people at a Planned Parenthood clinic in Colorado Springs, Colorado. In a post-arrest interview, the suspect is reported to have said "no more baby parts".

The episode has shone a spotlight on a little-discussed arm of biomedical research, raising the questions of why, how and how widely fetal tissue is used. To find out, *Nature* turned to an NIH database of research grants funded in 2014 to find those using fresh human fetal tissue, and in October contacted 18 researchers working with it. Su was one of only two who were willing to be interviewed. Most requests were declined or went unanswered; a public-affairs officer at one major Texas university refused to have a researcher speak to *Nature* to keep that person "safe".

The figures show that in 2014, the NIH funded 164 projects using the tissue, at a cost of \$76 million. This is slightly less than half of what the agency spent on work with human embryonic stem cells (ES cells), which has also been highly controversial, and 0.27% of the \$27.9 billion it spent on all research. (By comparison, the UK Medical Research Council spent 0.16% — £1.24 million (\$1.9 million) — of its total spending on research on five projects involving fetal tissue in the 12 months up to 31 March 2015.) Analysis of the NIH projects shows that the tissue is used most heavily for research on infectious diseases, especially HIV/AIDS; in the study of retinal function and disease; and in studies of normal and anomalous fetal development (see 'Fetal tissue research by discipline').

Opponents argue that the work is not necessary because other model systems and techniques can be used. "This is antiquated science," says David Prentice, the vice-president and research director at the Charlotte Lozier Institute, the research arm of the Susan B. Anthony List, which is an anti-abortion organization in Washington DC. "There are better and, frankly, more successful alternatives."

But supporters of the research counter that fetal tissue is legally obtained, that it would otherwise be destroyed, that such work has already led to major medical advances and that, if there were better alternatives, they would turn to them. "Fetal tissue is a flexible, less-differentiated tissue. It grows readily and adapts to new environments, allowing researchers to study basic biology or use it as a tool in a way that can't be replicated with adult tissue," says Carrie Wolinetz, the NIH's associate director for science policy.

"I get very frustrated when misinformed people go on about how it can all be done with computer models or cell cultures or stem cells or animals," says Paul Fowler, a reproductive biologist at the University of

Aberdeen Institute of Medical Sciences, UK, who in January published a study using livers from aborted fetuses to probe the impacts of maternal smoking on liver development<sup>1</sup>. "In some areas, the human is absolutely dramatically different than rodents."

Some argue that the entire episode represents a thinly cloaked attempt to attack and limit access to abortion by eroding support and funding for Planned Parenthood. "People are talking about fetal tissue, but really what this discussion is about is abortion," says Shari Gelber, a specialist in maternal-fetal medicine at Weill-Cornell Medical College in New York City, who has argued for the value of the research.

## LABORATORY LINES

Cell lines derived from aborted fetal tissue have been fairly commonplace in research and medicine since the creation in the 1960s of the WI-38 cell strain, which was derived at the Wistar Institute in Philadelphia, Pennsylvania, and MRC-5, which came from a Medical Research Council laboratory in London (see *Nature* **498**, 422–426; 2013). Viruses multiply readily in these cells, and they are used to manufacture many globally important vaccines, including those against measles, rubella, rabies, chicken pox, shingles and hepatitis A.

An estimated 5.8 billion people have received vaccines made with these two cell lines which, with others, have become standard laboratory tools in studies of ageing and drug toxicity. (Research with such lines is not covered by US regulations governing the use of fresh fetal cells and tissue nor captured in the NIH database.) In the past 25 years, fetal cell lines have been used in a roster of medical advances, including the production of a blockbuster arthritis drug and therapeutic proteins that fight cystic fibrosis and haemophilia.

But off-the-shelf fetal cell lines are of limited use for scientists because they do not faithfully mimic native tissue and represent only a subset of cell types: WI-38 and MRC-5, for example, were derived from fetal lungs. The lines can also accumulate mutations after replicating *in vitro* over time. And creating humanized mice such as Su's requires whole pieces of fetal organs to provide sufficient numbers of stem cells. For all of these reasons, researchers turn to fresh tissue.

In the United States, this is collected at medical centres and clinics that perform abortions under a patchwork of laws and regulations governing consent, tissue collection and transfer (see 'Fetal tissue and the law'). US law says that clinics can recover "reasonable payments" to offset the costs of providing the tissue, but it makes it a felony to profit from doing so. Planned Parenthood officials say that its clinics obtain full and informed consent from women choosing to donate fetal remains for research, and the organization

announced in October that its clinics will no longer recover costs of \$45–60 per specimen for collecting the tissue.

From the clinics, fetal tissue is then often passed to biological-research supply companies, which act as intermediaries and process the tissue before selling it to researchers. Su pays \$830 for each sample of fetal liver tissue supplied to his lab by one of the most widely used suppliers, Advanced Bioscience Resources in Alameda, California.

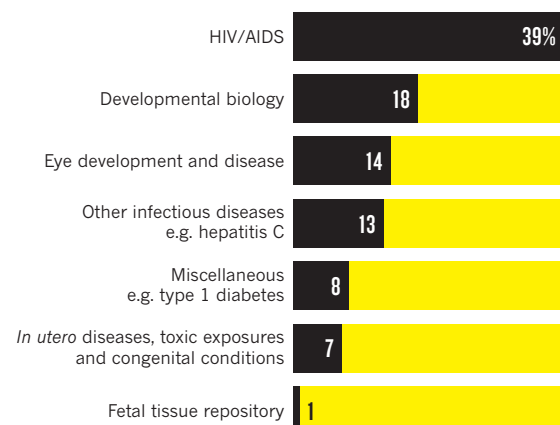
## HIV AND AIDS

The category of fetal tissue work that draws most NIH funding is the study of HIV and AIDS: it accounts for 64 of the 164 NIH grants. Researchers in this field have long struggled with the paucity of effective models for this uniquely human disease. The standard models, macaques, are expensive to breed, are infected with SIV instead of HIV and have immune responses that are different from those of people. The flexibility and adaptability of fetal tissue — and its richness as a

"USING FETAL TISSUE IS NOT AN EASY CHOICE, BUT SO FAR THERE IS NO BETTER CHOICE."

## FETAL TISSUE RESEARCH BY DISCIPLINE

The US National Institutes of Health funded 164 projects using human fetal tissue in the 2014 fiscal year, in these research areas:



## US REGULATION

### *Fetal tissue and the law*

Regulations governing US-funded fetal tissue research, first issued in 1975, state that:

- The research must comply with all applicable US, state and local laws and regulations.
- If information associated with the fetal tissue allows it to be traced to a living individual, that person becomes a research subject and informed consent from the donor is required for its use.

(Laws in at least 40 states require informed consent from the woman even if the fetal tissue will be anonymized.)

Additional requirements from a 1993 US law:

- Providers may not transfer fetal tissue for profit, but can receive funds to cover 'reasonable payments', such as for processing, storage and transportation.
- Researchers may not acquire fetal tissue if they know that a pregnancy was initiated in order to provide that tissue for research.
- Violators of either provision above are subject to criminal penalties of up to ten years in prison, up to US\$500,000 in fines, or both. These apply to both the tissue supplier and the tissue receiver in a transaction.

source of stem cells — has allowed the creation of a number of mice with humanized immune systems.

Prominent among these is the BLT (bone marrow–liver–thymus) mouse, which was created in 2006 (ref. 2). This model is made by destroying the animal's immune system and then surgically transplanting liver and thymus tissue fragments from a human fetus into the mouse. The immune system is further humanized with a bone-marrow transplant, using blood-forming stem cells from the same fetal liver. The animal enables studies of, for instance, immune responses that are key to developing an effective HIV vaccine. The mouse has "accelerated the study of HIV pathogenesis and novel approaches to harness anti-viral immunity to control HIV", reads a recent review by several

NIH-funded scientists who are using the mouse<sup>3</sup>.

The mouse has also helped to demonstrate that prophylactic drugs may prevent vaginal HIV infection — a strategy that is now in late-stage human trials. The animal is currently being used to examine how genital infection with herpes simplex virus alters immunity at the vaginal mucosa, making it easier for HIV to infect. In a similar vein, Su is now using his humanized mouse to examine the mechanisms by which hepatitis C and HIV co-infection can hasten liver disease.

There are drawbacks: the BLT mouse's average lifespan is relatively short, at only around 8.5 months, because the animals tend to develop cancers of the thymus. And the humanized immune system is not inherited, so the model must be created again and again — leading to the constant demand for fetal tissue that so disturbs abortion opponents.

## HUMAN DEVELOPMENT

In some research areas, fetal tissue may, in time, be replaced by other materials and methods: alternative, flexible cell types, including human ES cells and induced pluripotent stem (iPS) cells, and organoids, which are lab-created cellular structures that resemble tissue from normal organs (see *Nature* 523, 520–522; 2015). But there is one area in which, scientists say, fetal tissue is needed by definition: studies of early human development, and why it sometimes goes wrong.

"Human fetal tissue is likely never going to be replaced in some areas of research, particularly relative to fetal development," says Wolinetz. And the application of such work goes far beyond understanding developmental disorders such as congenital heart disease or other malformations, says Neil Hanley, an endocrinologist at the University of Manchester, UK. "For a wide range, now, of adult diseases and disorders, we know that they have their origins during very early human development," he says — type 2 diabetes and schizophrenia are both cases in point. "And unless you understand normal you're not going to understand abnormal."

The 30 developmental-biology grants involving fetal tissue that were awarded by the NIH in 2014 range from a study of the differentiation of myoblasts, which are the embryonic precursors of muscle cells, to several examinations of development of the urogenital tract — studies with relevance, for instance, to hypospadias, a common condition in which the urethra fails to close and the underside of the penis is incompletely formed. One project is creating a three-dimensional atlas of gene expression in the genital tubercle, the precursor of the penis. Another is probing gene activity in cells lining the fetal intestine to help explain excessive intestinal inflammation in premature babies. Hanley says that such studies are important, particularly because gene regulation — the finely tuned symphony that controls when and where genes are active — can vary strikingly between species, so findings in other animals often do not hold true in humans.

More than half of the 30 grants are for studies of brain development, and many of these projects are seeking advances in combating maladies such as autism, schizophrenia and Alzheimer's disease. Larry Goldstein, a neurobiologist at the University of California San Diego School of Medicine in La Jolla, uses cells called astrocytes from the brains of aborted fetuses to nourish neurons that he has derived from iPS cells and that have mutations associated with Alzheimer's disease. The astrocytes are thought to secrete factors that keep the neurons healthy in culture, and he uses the system to study the pathogenesis of the disease and to test potential drugs.

Goldstein hopes eventually to derive the astrocytes, too, from iPS cells. But "the human fetal astrocytes that we get at present are the gold standard that we use, and will use, to compare astrocytes that we make by differentiation", he says. He has also used neurons from aborted fetal brains to compare with the neurons made from iPS cells<sup>4</sup>. "As long as fetal tissue is available, this is a very valuable use of it," he says.

Another 23 of the NIH grants using fetal tissue involve eye development and disease. Damage to the retinal pigment epithelium (RPE), a single layer of cells at the back of the eye, has a key role in a





The collection of aborted fetal tissue for use in research has prompted demonstrations for and against US health provider Planned Parenthood.

number of eye diseases, including age-related macular degeneration, the most common cause of blindness in adults in the developed world. The 2000s saw advances in ways to create cell cultures with RPE dissected from the eyes of fetuses, allowing scientists to study the function of these cells in a dish. And although some scientists have turned to stem cells to generate RPE, like Goldstein they continue to use fetal tissue as a benchmark of normal development and function.

Goldstein agreed to speak to *Nature*, he says, because “somebody has to speak up responsibly”. He stressed that he and his colleagues think hard about the ethics of their work. “We are not happy about how the material became available, but we would not be willing to see it wasted and just thrown away.”

Occasionally, fetal tissue is used for clinical work. Last year, a company called Neuralstem in Germantown, Maryland, in collaboration with scientists at the University of California, San Diego, launched a trial in which stem cells from fetal spinal cord were implanted to treat spinal-cord injuries. In May, researchers in the United Kingdom and Sweden launched a study in which dopaminergic neurons from aborted fetuses are transplanted into the brains of patients with Parkinson’s disease (see *Nature* **510**, 195–196; 2014). Research with fetal tissue is less controversial in countries where abortion is more widely accepted.

### UNCOMFORTABLE VIEWING

The Planned Parenthood videos caused even some supporters of fetal tissue research to feel uncomfortable. In one video, physician Deborah Nucatola, the group’s senior director of medical services, describes how she crushes fetuses above and below key organs to preserve them intact for research. She also described turning a fetus into a breech presentation to deliver the head last, when the cervix is more dilated, thus preserving the brain.

This raised the question of whether physicians are altering abortion techniques to accommodate research requests, violating a widely held precept of research ethics. Arthur Caplan, a bioethicist at the New York University School of Medicine, dismisses the videos as “pure politics”, but some of the footage “did get my eyebrow to arch”, he says. “You can’t use a different approach to the abortion to try to preserve something. Those are just no-no’s.”

Planned Parenthood spokeswoman Amanda Harrington says that the organization is not aware of any instances in which the method of an abortion has been changed to preserve organs. But, she adds, “if minor

adjustments that have no bearing on the woman’s health and safety are done when the woman has expressed a desire to donate tissue, that is entirely appropriate and ethical and legal”. Women’s health and safety, she says, “is always the number one priority”.

The question for many scientists is what the fallout of the controversy will be. On the heels of the Colorado shootings, some Republicans in Congress backed off earlier attempts to defund Planned Parenthood, and President Obama is expected to veto any bill that does so. This means that the lasting damage of the videos may end up being inflicted not on Planned Parenthood’s budget, but on science. Since July, four bills that would criminalize or otherwise restrict the research have been introduced in the US Congress, and lawmakers have launched similar efforts in a dozen state legislatures. (Missouri, Arizona and North Dakota already ban the research.)

Su felt the climate for his research grow colder when, on 1 October, a new North Carolina law was signed that makes it a felony to sell fetal tissue for any amount within the state. Su receives the tissue he uses from outside the state, but the message behind the new law concerns him. “I hope this current controversy, or possible congressional interventions, won’t slow down biomedical research,” he says. “The benefit is bigger than the drawback on this.”

The controversy “absolutely puts fetal tissue research at risk”, says Caplan. “Young scientists are unlikely to enter a field riven with controversy, where funding is uncertain and physical threats are a real possibility.”

Caplan says that parallels could emerge with events in the early 2000s, when the use of human ES cells in US research became politically fraught. Then, tight federal regulations governing NIH funding of the research were adopted, but some states, including California and Massachusetts, responded by pouring money into the science all the same.

“To move ahead, the reality is that fetal tissue research need not be funded or permitted everywhere,” Caplan says. “It needs to be allowed somewhere.” ■

**Meredith Wadman** is a freelance writer based in Virginia and an editorial fellow at New America, a think tank in Washington DC.

1. Drake, A. J. *et al.* *BMC Med.* **13**, 18 (2015).
2. Melkus, M. W. *et al.* *Nature Med.* **12**, 1316–1322 (2006).
3. Karpel, M. E., Boutwell, C. L. & Allen, T. M. *Curr. Opin. Virol.* **13**, 75–80 (2015).
4. Israel, M. A. *et al.* *Nature* **482**, 216–220 (2012).



# THE LAST ILLNESS

RESEARCHERS ARE  
GAINING INSIGHT  
INTO THE CAUSES  
OF CACHEXIA —  
A DEVASTATING FORM  
OF MUSCLE WASTING  
THAT IS OFTEN THE  
FINAL STAGE OF  
CANCER AND  
OTHER DISEASES.

BY CORIE LOK

Cachexia is a familiar condition in palliative-care departments such as this one in Puteaux, France.

**A**s a palliative-care researcher, Susan McClement has talked to many people dying of cancer and their families — and some of their stories are burned into her brain. One man was so concerned by the sight of his emaciated wife, whose body had been ravaged by metastatic breast cancer, that he resorted to force feeding her — pinching her nose and slipping in a spoonful of food when she opened her mouth. Convinced that food would give her the energy to fight the cancer, his daily visits became protracted battles. She died a few weeks later.

McClement, who works at the University of Manitoba in Winnipeg, Canada, says that nutritional conflicts can become a source of regret for relatives. “They said, ‘You know, if I could do it over again, I would have spent much less time fighting about tapioca pudding and much more time telling my wife that I loved her.’”

The woman in this case had cachexia, a metabolic disorder that affects some 9 million people worldwide, including as many as 80% of people with advanced cancer. It typically involves extreme weight- and muscle-loss, makes routine activities difficult and increases the risk of deadly complications such as infections. Adding calories doesn’t reverse cachexia, and McClement says that the disorder sometimes provokes extreme reactions from family members because it serves as visual confirmation of their worst fears. “It’s a constant reminder that the person is sick and is not going to get better,” says McClement.

Cachexia is seen in the late stages of almost every major chronic illness, affecting 16–42% of people with heart failure, 30% of those with chronic obstructive pulmonary disease and up to 60% of people with kidney disease. But for

many years it was overlooked, as physicians and researchers focused their attention on the primary illness instead.

Now, scientists are increasingly viewing cachexia as a distinct, treatable condition. Basic research has revealed how it is driven by inflammation and metabolic imbalances, and has generated drug targets, says Stefan Anker, a cardiologist and cachexia specialist at the University Medical Center Göttingen in Germany. “Now we have quite a number of powerful options to test,” he says. This has spurred investment from drug developers who aim to reduce suffering, and possibly give patients the strength to withstand chemotherapy or surgery.

But some high-profile clinical trials in the past two years have produced disappointing results, prompting much self-reflection in the young field. “I’m a little bit worried that if we

VOISIN/PHANIE/CORBIS



don't see a successful clinical trial in the next five years, the dollars from the pharmaceutical industry to develop a treatment will go somewhere else," says Jose Garcia, a clinical researcher focused on wasting disorders at the Michael E. DeBakey Veterans Affairs Medical Center in Houston, Texas. "In my view, that would be a missed opportunity."

### WASTED ENERGY

The term cachexia is derived from the Greek *kakos* and *hexis*, meaning 'bad condition'. It is thought that Hippocrates recognized the syndrome — but it took until 2006 for the cachexia field to start working up a formal definition, which includes a loss of 5% or more of body weight over 12 months, and reduced muscle strength. In the clinic, it remains under-recognized by oncologists, says Egidio Del Fabbro, a palliative-care physician and researcher at Virginia Commonwealth University in Richmond. There are no standard guidelines for treatment.

In the past decade, researchers have made strides in learning about the causes of cachexia, thanks to funding from the US National Cancer Institute and some advocacy groups. New international conferences (including one that wrapped up this week in Paris) and the launch of a research journal — the *Journal of Cachexia, Sarcopenia and Muscle* — have also drummed up interest in the field.

It is now clear that a key mechanism underlying cachexia is the increased breakdown of muscle protein, along with dampened protein synthesis, which leads to overall muscle loss. Studies in 2001 helped to jump-start the field when they identified genes that were more active in atrophying rodent muscles than in normal ones<sup>1,2</sup>. These genes encode enzymes called E3 ubiquitin ligases, which tag proteins for destruction in the cell. Mice without these enzymes were resistant to muscle loss.

Muscle cells seem to make more of these ligases when hit with certain inflammatory signals from tumours or from immune cells responding to cancer or other illness. Abnormalities in apoptosis (programmed cell death) and in the muscle cell's energy-producing organelles, mitochondria, have also been implicated.

Several drug-makers have homed in on the protein myostatin, which blocks muscle growth. In a 2010 paper<sup>3</sup> that got many people excited about a possible cachexia drug, researchers from biotechnology company Amgen in Thousand Oaks, California, showed that they could reverse muscle loss and extend the lives of mice with tumours and cachexia by blocking signalling through the myostatin pathway.

Research since then suggests that cachexia is more than a muscle disease. Studies<sup>4</sup> have identified problems in the brain's regulation of appetite and feeding, and even ways in which the liver might be contributing to the energy imbalance that sees the body burn its own tissue to sustain itself. Others have looked at fat tissue, which can also waste away in cachexia.

They showed that inflammation<sup>5</sup> and molecules made by tumours<sup>6</sup> cause white fat cells to turn into brown fat cells, which burn more energy to generate heat than white fat cells. The question that researchers are now tackling is how tissues and organs — muscle, brain, fat, even bone — are communicating with one another. A paper published last week<sup>7</sup> suggests that fat signalling could be involved in muscle atrophy.

All this research has brought more representatives of biotechnology and pharmaceutical companies to cachexia meetings in recent

## "IT'S A CONSTANT REMINDER THAT THE PERSON IS SICK AND IS NOT GOING TO GET BETTER."

years, says Denis Guttridge, a cell biologist at the Ohio State University in Columbus, who organizes one such conference. "That's exciting for a basic scientist like myself," he says. "I can see the increase in the translational pipeline."

### DRUG DISAPPOINTMENT

Despite the excitement in labs, clinical research has so far proved disappointing. In 2011, biotech firm GTx of Memphis, Tennessee, launched two late-stage clinical trials of enobosarm, a molecule that binds to the same receptor as testosterone but only in muscle and bone, mimicking the hormone's ability to stimulate muscle build-up but without its undesirable side effects. Results from earlier, smaller trials looked promising: people taking the drug had increased lean body mass and improved physical function, as measured by their speed at climbing stairs<sup>8</sup>. But in the larger tests of the drug, on people with advanced lung cancer, the benefits in function disappeared. The firm has since abandoned muscle wasting, and is instead testing larger doses of enobosarm to treat breast cancer.

A pair of unpublished studies on people with lung cancer and cachexia tested a compound called anamorelin, which mimics ghrelin, an appetite-stimulating peptide hormone produced mainly by the stomach. The trials were sponsored by pharmaceutical company Helsinn in Lugano, Switzerland, which reported that participants in the treatment group put on weight and muscle mass compared with those taking a placebo, but showed no difference in hand grip strength. Still, the company announced last week that the European Medicines Agency is reviewing its drug for approval.

There is a lot of debate about why the trials failed to show functional improvements. Some researchers say that the teams did not use the most clinically relevant measures of muscle function. "We don't really know what is the best test for this," says Garcia. "If you can climb up a set of stairs one second faster, what does that mean?" This confusion about trial design is a problem for the field, says Anker. "We need to reach consensus on endpoints and what to aim for in our treatments."

Another problem is that animal data on cachexia may not translate into humans. Some work has tried to make a case that the mechanisms found in rodents might be similar to those in humans, by looking at human tissue samples, says Vickie Baracos, a clinical translational researcher in muscle wasting at the University of Alberta in Edmonton, Canada. "But held up to scrutiny, this clinical evidence is often rather sketchy."

Researchers in the field lament the dearth of human data and clinical samples. Baracos

says that studies are needed that follow people with cachexia over time, collecting blood and muscle samples along the way. "A cachexia data repository with a biobank would sure be a great thing," she says.

Perhaps the biggest challenge is that the field has to compete for funding and recognition with research into other major diseases, says Anker. "Cachexia is competing for internal resources within big companies, fighting with cancer, cardiology," he says. Few companies have dedicated cachexia groups or departments. GTx stopped its work on muscle wasting in part because insurers did not seem interested in covering a medication that was only going to target cachexia and not cancer, says Mary Ann Johnston, the company's vice-president for clinical development. "There's a lack of interest in supportive care."

But an effective treatment would be transformative, says Garcia. It might spur physicians to talk more to patients and their families about the troubling symptoms of cachexia. Without the tools to treat the syndrome, many doctors don't address it, he says. And that vacuum of information can be distressing.

McClement, for her part, has been interviewing more families of people with cachexia. She hopes to find ways to better inform them about the condition and help them to cope. Given the absence of pharmacological interventions, such psychosocial ones are important, she says. "That's all we've got." ■

**Corie Lok** is Nature's Research Highlights editor based in Cambridge, Massachusetts.

1. Bodine, S. C. *et al. Science* **294**, 1704–1708 (2001).
2. Gomes, M. D., Lecker, S. H., Jagoe, R. T., Navon, A. & Goldberg, A. L. *Proc. Natl Acad. Sci. USA* **98**, 14440–14445 (2001).
3. Zhou, X. *et al. Cell* **142**, 531–543 (2010).
4. Argilés, J. M., Busquets, S., Stemmler, B. & López-Soriano, F. J. *Nature Rev. Cancer* **14**, 754–762 (2014).
5. Petruzzelli, M. *et al. Cell Metab.* **20**, 433–447 (2014).
6. Kir, S. *et al. Nature* **513**, 100–104 (2014).
7. Kir, S. *et al. Cell Metab.* <http://dx.doi.org/10.1016/j.cmet.2015.11.003> (2016).
8. Dalton, J. T. *et al. J. Cachexia Sarcopenia Muscle* **2**, 153–161 (2011).

# COMMENT

**DRUGS** Push research and regulation for ionic liquids **p.188**

**MATHEMATICS** Exhaustive biography of Enlightenment colossus, Leonhard Euler **p.190**

**EVOLUTION** Eleven testers compare three board games about natural selection **p.192**

**CONSERVATION** Analysis questions EU funding priorities **p.193**



JESSICA SCRANTON



The impact of sustainability measures such as Mexico's Encrucijada Biosphere Reserve can be assessed by aggregating existing reviews and impact evaluations.

## Map the evidence

Too many studies go unread. Collate them to enable synthesis and guide decision-making in sustainability, urge **Madeleine C. McKinnon** and colleagues.

“What if someone had already figured out the answers to the world’s most pressing policy problems, but those solutions were buried deep in a PDF, somewhere nobody will ever read them?” asked a *Washington Post* blog last year.

It was on to something. Many of the tens of thousands of documents that are produced every year to assess the impacts of sustainability policies and programmes are never read<sup>1</sup>. In 2014, the World Bank<sup>2</sup>

found that almost one-third of its archived policy reports — documenting the impacts of its numerous projects, from dam construction to microcrediting — has never been downloaded.

It doesn’t have to be this way. Experts in evidence synthesis, a field that involves the use of various tools and methods to locate and combine many sources of data, are starting to produce evidence maps for wayfaring researchers and policymakers. These pull together and categorize systematic reviews,

impact evaluations and other primary-research studies in a particular area (such as agriculture or education), and visually distil the scope and effects of interventions that have been implemented<sup>3</sup>.

Evidence maps can show at a glance which areas or relationships have been studied most — whether it be the impact of ecotourism on local economies or of education on reducing harmful fishing practices. They can also highlight key gaps in the evidence base, and so guide the prioritization of research. ►



► We are an interdisciplinary working group supported by the Science for Nature and People (SNAP) Partnership that aims to understand how protecting nature can enhance human well-being (see [go.nature.com/fdsj4v](http://go.nature.com/fdsj4v)). We contend that evidence mapping should be applied to sustainable development much more broadly. To aid decision-making at the intersection of development and conservation, we have produced an interactive map that categorizes more than 1,000 studies documenting links between conservation efforts and human well-being. We urge policymakers and researchers working in sustainability to develop similar tools to enable researchers, donors and practitioners to rapidly find and assess the information relevant to them.

### MIND THE GAPS

Failure to evaluate the existing evidence can result in unnecessary harm. For example, conducting a systematic review in the 1960s of the available data on cot death (instead of waiting until the 1990s) could have led to earlier recognition of risk factors associated with infant sleeping positions and prevented an estimated 10,000 infant deaths in the United Kingdom alone<sup>4</sup>.

Another consequence may be overlooking the relative costs and benefits of different interventions. Development projects designed to prevent diarrhoeal disease, for instance, continue to emphasize the importance of providing people with access to clean water. Yet a 2012 synthesis<sup>5</sup> of different sanitation interventions indicates that prioritizing certain behavioural changes, such as hand-washing, can bring equivalent health and other benefits, and cost much less than introducing a new water supply.

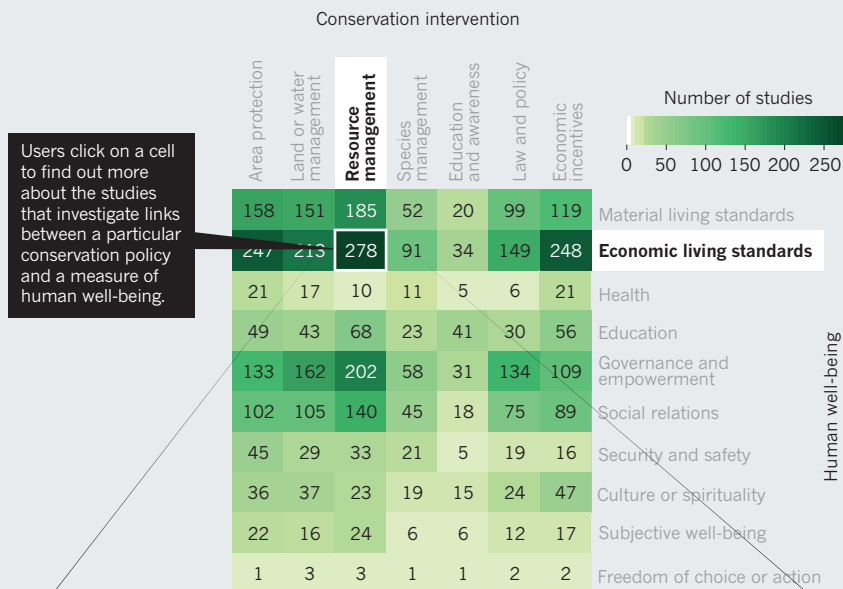
Often people assume that there is evidence to justify a particular intervention, or they act on the basis of what is familiar rather than what is proven. Establishing and maintaining national parks and other protected areas is one of the more prevalent conservation approaches used by governments and non-governmental organizations. Yet a 2013 systematic review<sup>6</sup> of qualitative and quantitative assessments of protected areas worldwide shows that many of the broad assumptions that underlie their creation — that protected areas have positive social impacts, say, or provide economic benefits through tourism — are not reliably supported.

Systematic reviews of multiple studies, such as the three just described, enable researchers to assess the quality of the available evidence and to make predictions about the effectiveness of a programme more broadly. Yet because documents are often inaccessible — hidden behind paywalls, or buried in hard drives and filing cabinets in field offices — obtaining the relevant publications and reports to conduct such reviews

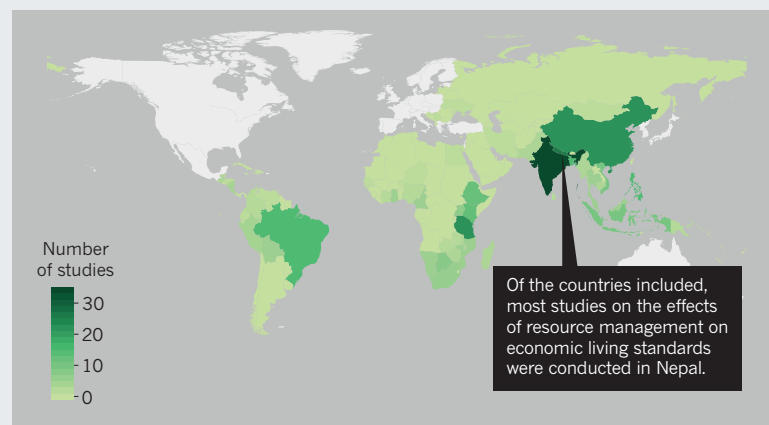
## NAVIGATING WHAT'S KNOWN

An evidence map is produced by collecting and categorizing studies that probe the link between conservation efforts and human well-being. It reveals where data are available to establish causal pathways, and where there are gaps in the knowledge.

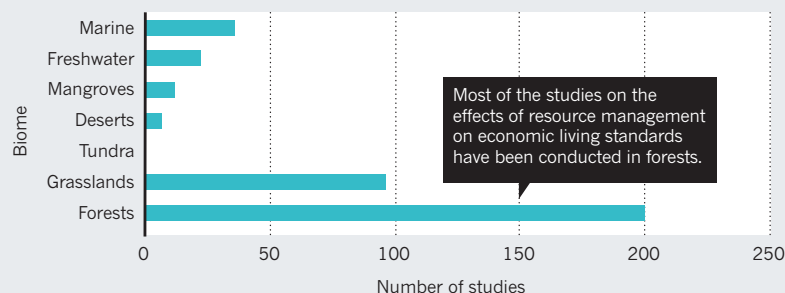
*More than 1,000 studies were included in the evidence map on the basis of systematically designed selection criteria.*



*A search of these 278 studies by country indicates where in the world they have been carried out.*



*Or these 278 studies can be searched by the biome in which they were conducted.*



is costly and time consuming.

The resources currently allocated to monitoring and evaluation — on average, less than 5% of a conservation project's budget — do not come close to what is needed to satisfy the increasing demand from policymakers for more and higher-quality evidence on the impacts of conservation and development interventions<sup>1</sup>.

## INTERNATIONAL EFFORTS

Researchers and others need new ways to prioritize efforts, to ensure that funded evaluations and systematic reviews address relevant questions and to track progress in the generation of evidence.

Encouragingly, various organizations have begun to produce and fund systematic reviews and high-quality evaluations to assess the effects of conservation interventions. Among them is the International Initiative for Impact Evaluation (3ie), a non-governmental organization that promotes evidence-informed policies and programmes for development.

Standards, tools and practical guidance — for instance, on steps to reduce bias — have also begun to emerge. For instance, international research networks, such as the Collaboration for Environmental Evidence, are promoting the use of rigorous methodologies to assess the environmental, social and health impacts of specific interventions<sup>2</sup>, such as planting hedgerows or the use of genetically modified crops. The Evidence for Policy and Practice Information and Co-ordinating Centre at the UCL Institute of Education is developing various tools to help researchers to retrieve, characterize and extract data when doing evidence synthesis. We think that evidence maps should be added to this growing stock of synthesis tools.

Over the past five years, 3ie, and researchers funded by the organization, have generated five maps for a broad range of topics, from education to sanitation and hygiene. These maps identify well-studied areas, such as the effects of hand-washing and other sanitation measures on the incidence of diarrhoea. The maps also flag gaps in research. For instance, little is known about the impact of education-related interventions on household expenditure or child labour. As far as we know, no other evidence maps relevant to sustainable development have been created.

We have extended 3ie's effort by creating a map of the scope and amount of existing evidence on the impact of nature conservation on human well-being globally<sup>3</sup>. To do this, we mined online databases and websites and asked individuals who were well-positioned to alert us to studies. We narrowed our search to studies published since 1970 that had been conducted in developing countries. We then took systematic steps — based

on factors such as study design and type of conservation intervention — to winnow our collection. Following 3ie, we included systematic reviews and impact evaluations. We also categorized other primary research studies, such as non-comparative studies in which no control situation was available to compare the causal effect of the intervention.

Our map reveals that around 25% of the 1,014 studies we categorized measured the economic impacts of establishing and maintaining protected areas, whereas only 2% considered equally important measures of well-being, such as the health of local communities (unpublished work). More data, for instance, on the country and biome in which the study was conducted and on the study design, enable users to explore in more detail the distribution and robustness of the evidence base (see 'Navigating what's known'). For instance, we were surprised to learn that since 1970, almost half of all rigorous impact evaluations have been conducted in only six countries — including Tanzania, China and Thailand. Eight of the 90 such studies were carried out in Costa Rica, probably in part thanks to the country's strong governance in environmental issues and its well-established research community.

Our map also exposes areas where more data are needed. In Indonesia, for example, little is known about the relationship between conserving marine ecosystems and the sense of empowerment felt by local communities. And it illuminates which relationships are well-studied and ripe for analysts to explore trends and examine causal pathways through full systematic reviews — such as those between fisheries resource management and economic and material well-being.

## GLOBAL CHALLENGES

Evidence maps are needed for a broad range of topics that are central to sustainable development, such as renewable energy, food security and disaster-risk management. As new evidence becomes available, such maps will need to be updated, ideally by a central curator, to enable researchers to track progress in plugging knowledge gaps. Publishers are well placed to catalyse the development of such shared map resources by enabling broader access to their journals and encouraging the use of standardized search terms and approaches in individual studies. Donors should make the placement of a study in an open evidence repository a condition of funding — much as public funding agencies, such as the US National

*“Since 1970, almost half of all rigorous impact evaluations have been conducted in only six countries.”*

Institutes of Health, require grant recipients to make their research openly available to the public through PubMed Central.

The Sustainable Development Goals (or Global Goals) were launched in September by the United Nations. Among the 17 priorities for governments, businesses and others are addressing climate change, alleviating poverty and malnutrition and protecting the planet's natural resources. Information on what kinds of interventions are effective in helping the world to meet these and other challenges, and under what conditions, will be crucial to guiding policy and tracking progress<sup>4</sup>.

An evidence atlas for sustainable development, progressively updated and improved, could help to transform the work of addressing global challenges into a rigorous science. ■

**Madeleine C. McKinnon** is senior director of monitoring and evaluation at Conservation International, Arlington, Virginia, USA. **Samantha H. Cheng** is a doctoral candidate at the University of California, Los Angeles, California, USA. **Ruth Garside** is senior lecturer in evidence synthesis at the European Centre for Environment and Human Health, University of Exeter Medical School, Truro, UK. **Yuta J. Masuda** is an environmental and social-policy scientist at the Nature Conservancy, Boston, Massachusetts, USA. **Daniel C. Miller** is assistant professor in natural resources and environmental sciences at the University of Illinois, Urbana, Illinois, USA.  
e-mail: mmckinnon@conservation.org

1. Fisher, B. *et al.* *Conserv. Biol.* **28**, 880–882 (2014).
2. Doemeland, D. & Trevino, J. *Which World Bank Reports are Widely Read?* (World Bank, 2014).
3. Snilstveit, B., Vojtkova, M., Bhavsar, A. & Gaarder, M. *Evidence Gap Maps: A Tool for Promoting Evidence-Informed Policy and Prioritizing Future Research* (World Bank, 2013).
4. Gilbert, R., Salanti, G., Harden, M. & See, S. *Int. J. Epidemiol.* **34**, 874–887 (2005).
5. Rijsberman, F. & Zwane, A. P. *Water & Sanitation* (Copenhagen Consensus, 2012).
6. Pullin, A. S. *et al.* *Environ. Evid.* **2**, 19 (2013).
7. Collaboration for Environmental Evidence. *Guidelines for Systematic Review and Evidence Synthesis in Environmental Management Version 4.2* (Environmental Evidence, 2013).
8. Bottrill, M. *et al.* *Environ. Evid.* **3**, 16 (2014).
9. Lu, Y., Nakicenovic, N., Visbeck, M. & Stevance, A. S. *Nature* **520**, 432–433 (2015).

For a list of co-signatories to this article, see [go.nature.com/qiuwlk](http://go.nature.com/qiuwlk).

## CORRECTION

The Comment article 'Find asteroids to get to Mars' (R. P. Binzel *Nature* **514**, 559–561; 2014) omitted sources for the graphic 'Mission requirements'. The credit has been updated online at [go.nature.com/6je8v3](http://go.nature.com/6je8v3).





# Develop ionic liquid drugs

Update regulation to spur research into drugs that the body absorbs more easily and that could reach market more quickly, urge **Julia L. Shamshina** and colleagues.

**A** new realm of potential drugs is hiding in plain sight. Pharmaceutical research, manufacture and regulation focuses on solid active ingredients, delivered as powders or tablets. Liquid forms are neglected and viewed as an intermediate step, rather than an endpoint.

Yet many promising solid drug candidates are too insoluble for the body to absorb. Of

the compounds entering development, 40–70% fail because they cannot be modified simply to allow effective release into the bloodstream<sup>1</sup>.

Meanwhile, ionic liquids, an exciting class of chemical that could bypass these delivery problems, are being ignored<sup>2</sup>. Half of all drugs sold are salts<sup>3</sup> that are held together by ionic bonds, among other forces. Salts that

are liquid at room or body temperature can have dramatically better solubility, absorba-

bility and stability than do solid forms<sup>4</sup>. Ionic liquids can also be configured to deliver two or more active ingredients at once.

For example, combining active ions from the pain reliever procaine and the non-steroidal anti-inflammatory drug (NSAID) salicylic acid generates a liquid salt, procainium salicylate. It could deliver the medical benefits of both compounds more efficiently and cheaply while opening up new treatment options.

With the drug-discovery pipeline clogged, it is time to try alternatives. We call on chemists and the pharmaceutical industry to develop liquid salt forms of drugs. Chemists will need to learn more about the spectrum of interactions in ionic liquids, how to engineer ionic bonds, and how the choice of ions changes the chemical, physical and biological properties of ionic compounds. Regulations must be updated to consider active ingredients in liquid as well as solid states.

## OLD HABITS

Why are ionic liquids being ignored? First, most academic and industrial chemists lack understanding and experience of working with them. Chemistry courses and textbooks teach that new molecules are made by manipulating covalent bonds (where electrons are shared between atoms) rather than ionic ones.

Second, pharmaceutical companies are conservative. Ionic liquids are unfamiliar, unregulated and felt to be too risky to develop commercially.

And there is a perception problem. Over the past 20 years many researchers (including us) have demonstrated the value of ionic liquids as solvents, electrolytes and compressor fluids that are reusable, non-volatile and safe. Yet many researchers and journalists still associate the term with the first such chemicals widely studied: for example, dialkylimidazolium, quaternary ammonium and phosphonium salts, which were explored in the 1990s as potential 'green' solvents and electrolytes. With each new study, the class as a whole became pigeonholed — as expensive, cheap, green, toxic, biodegradable, non-biodegradable, non-flammable, flammable, volatile or non-volatile. In reality, ionic liquids have an infinite range of characteristics.

The effect of an ionic liquid on a living organism (such as toxicity) should be exploited rather than being seen as a problem, especially given that liquid salts can be more soluble in the body than are solids. The rate at which a drug is taken up depends on many factors, including its solubility, permeability, dissolution rate and the metabolic pathways involved.

The solubility of a solid depends so much on its form — for example, the size of the particles and whether it is crystalline or amorphous — that the same dose could be ineffective or toxic in different forms. That form depends unpredictably on the conditions under which the drug is manufactured and stored — temperatures, heating and cooling rates and solvents used<sup>5</sup>. In the late 1990s, for example, production of the antiretroviral ritonavir was halted temporarily when it emerged that an undiscovered, less soluble crystal structure formed from the manufactured drug capsules, making the medication ineffective. Laborious screening of solid forms is now part of the gauntlet of development stages for drugs — steps that would not be necessary for liquid pharmaceuticals.

Ionic liquids will not of course be problem-free. They may absorb contaminants, leach out of carriers, or be prone to subtle differences in biological activity that we do not yet know about. But, in our view, their huge potential for improving drug development, delivery and efficacy is ripe for exploration.

## SCIENTIFIC UNDERSTANDING

There is still much to learn. Academic researchers are starting to study the subtleties of ionic bonding in the liquid state. Swapping ions can render a salt soluble in or immiscible with a given solvent, stable or reactive, non-volatile or distillable, and permeable or not. There are many unknowns: how can we predict whether a given ion can be made liquid? How will we purify active ingredients that never crystallize?

Fine distinctions in ionic bonds may be controlled. Chemists are investigating how proton transfer (the moving of a hydrogen ( $H^+$ ) ion between an acid and a base) and hydrogen bonding (an electrostatic attraction between a hydrogen atom bound in a molecule and another nearby atom) influence the melting points of salts, how easily ions are transported across cell membranes, and the creation of different types of acid-base complexes<sup>6,7</sup>.

Industry is starting to take notice. A patch for treating lower back pain based on an ionic liquid combination of the NSAID etodolac and the pain-reliever lidocaine has finished phase III clinical trials (see [go.nature.com/urzlks](http://go.nature.com/urzlks)); the next step is to bring it to market. Tests of more ionic liquids and drug combinations are needed, with a wider variety of forms. For instance, an ionic liquid drug loaded on a powder could be taken orally, with the drug's dissolution rate tuned by changing the properties of the powder carrier.

Laboratory studies show how drug properties influence their flow around the body. We have shown, for example, that a liquid combination of lidocaine (a base) and ibuprofen (an

acid) forms strong hydrogen bonds between the two components that allow both to pass together through a cell membrane<sup>7</sup>. Likewise, the penetration of other drugs might be enhanced. Whether or not hydrogen-bonded complexes stay together or break apart in the bloodstream is already a topic of debate in the design of solid pharmaceutical co-crystals (crystalline blends of compounds)<sup>8</sup>.

## REGULATORY GAPS

For industry, the most important question is how ionic liquid pharmaceuticals will be defined and regulated. The US Food and Drug Administration (FDA) guidelines for active pharmaceutical ingredients (see [go.nature.com/ivpakv](http://go.nature.com/ivpakv)) focus on pure compounds and their stability, and solids are easier to classify and study. The latest guidelines, from April 2013, distinguish crystalline hydrogen-bonded co-crystals and ionized salts — which may differ only in their bonding — even though there is no recognized difference in their biological effects<sup>8</sup>.

The FDA considers a co-crystal a blend of a known active pharmaceutical ingredient and inert filler, which does not require stringent tests. It classes salts, on the other hand, as new compounds that require full-length investigation. Physically, such a distinction could correspond to as little as a 0.1-ångström shift of a hydrogen atom position in a crystal.

Ionic liquids do not fit these regulatory boxes. Different solid forms can be distinguished through measurements (such as X-ray diffraction or melting points). Such methods cannot detect the subtle difference in a liquid between an ionized salt and a hydrogen-bonded moiety. Protons move around in a liquid and may be in any state relative to a nearby ion<sup>9</sup>. Whether liquid compounds are ionized, neutral or hydrogen bonded also depends on temperature and the ingredients involved. Whereas solids retain their identity when they are mixed, blends of free ions in a liquid interact in myriad ways.

## NEXT STEPS

Rather than abandon the field as too boulder-strewn, as some chemists have suggested, researchers should define and understand ionic liquids so that regulation can keep pace with the discovery of new therapeutics.

Chemists in academia and industry need to find techniques for proving the structure and purity of ionic liquids. Methods must be developed for identifying certain complexes, free ions or dissociated acids and bases. Classification systems that encompass the varying degrees of ionic bonding should

be defined, and methods for assessing the purity of ionic liquids need to be developed. For clinical translation, delivery mechanisms and steps such as demonstration of shelf life will be needed.

Regulators should focus less on the form of a material and more on its clinical properties. Developing a scientific classification system for ionic liquids based on how they perform as drugs will enable regulatory agencies to make legal distinctions in a medical context. Distinctions will need to be made, for instance, between salts that dissociate in solution (and essentially behave like their solid counterparts in the bloodstream) and those that remain associated.

Many of the problems facing ionic liquid pharmaceuticals are shared with nanotechnology: toxicity, relative expense and a need for new terminology. For nanotechnology, investment made those tractable; the pharmaceutical industry should likewise invest.

Designing ionic liquids based on pharmaceuticals already approved by the FDA is a good place to start. Familiarity and success will encourage further investment. There is nothing magic in the design and study of a liquid, it is simply a matter of spending time and money to gain the understanding needed to overcome perceived chemical, biological, manufacturing and regulatory difficulties.

Given its importance to society, the pharmaceutical industry must open its mind to new approaches. ■

**Julia L. Shamshina** is chief technology officer at 525 Solutions in Tuscaloosa, Alabama, USA. **Steven P. Kelley** is a research assistant in the Department of Chemistry, University of Alabama, Tuscaloosa, Alabama, USA. **Gabriela Gurau** is an academic associate in green chemistry at McGill University, Montreal, Canada, and chief financial officer at 525 Solutions. **Robin D. Rogers** is a Canada Excellence Research Chair in green chemistry and green chemicals at McGill University, Montreal, Canada. e-mail: [robin.rogers@mcgill.ca](mailto:robin.rogers@mcgill.ca)

The authors declare conflicting financial interests, see [go.nature.com/phinz5](http://go.nature.com/phinz5) for details.

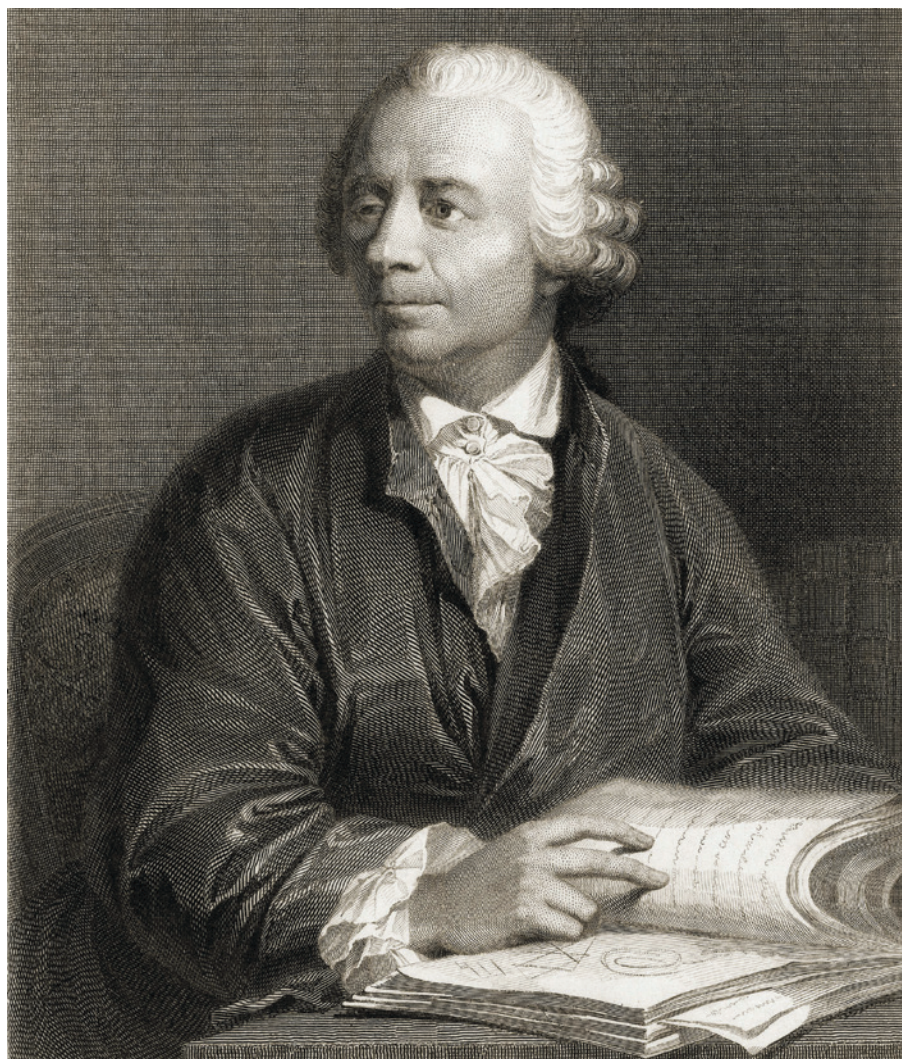
1. Hauss, D. J. *Adv. Drug Deliv. Rev.* **59**, 667–676 (2007).
2. Stoimenovski, J., MacFarlane, D. R., Bica, K. & Rogers, R. D. *Pharm. Res.* **27**, 521–526 (2010).
3. Stahl, P. H. & Wermuth, C. G. (eds) *Pharmaceutical Salts: Properties, Selection, and Use* 2nd edn (Wiley-VCH, 2008).
4. Shadid, M. et al. *MedChemComm* **6**, 1837–1841 (2015).
5. Stahly, G. P. *Cryst. Growth Des.* **7**, 1007–1026 (2007).
6. Shamshina, J. L., Barber, P. S. & Rogers, R. D. *Expert Opin. Drug Deliv.* **10**, 1367–1381 (2013).
7. Wang, H. et al. *Chem. Sci.* **5**, 3449–3456 (2014).
8. Aitpamula et al. *Cryst. Growth. Des.* **12**, 2147–2152 (2012).
9. Kelley S. P. et al. *Cryst. Growth Des.* **13**, 965–975 (2013).



## HISTORY

# A mathematical revolutionary

Davide Castelvecchi reviews a hefty biography of the prolific Enlightenment luminary Leonhard Euler.



Leonhard Euler's analysis gave physics and astronomy their modern mathematical shape.

The Swiss mathematician Leonhard Euler (1707–83) was all but blind when he moved to St Petersburg in 1766 for a second stint as the star of the Russian Imperial Academy of Sciences. He had lost vision in his right eye 28 years before; a cataract was claiming his left. Yet Euler, then 59, boasted in one letter that the loss of sight meant “one more distraction removed”. From then on, his productivity increased: more than half of his eventual output of 866 publications was

published either over the remaining 17 years of his life or posthumously.

The publication of Euler's collected works, which began in 1911, is still unfinished; it will fill more than 80 large volumes. There are also hundreds of letters — many with as much scientific content as the best of his papers. *Leonhard Euler*, written by historian of mathematics Ronald Calinger, is perhaps the first biography that attempts to offer a panoramic view of this immense body of work.

Euler dominated almost all branches of mathematics, as well as physics, astronomy and engineering, during the Enlightenment era. Euler's mathematics was often ahead of his time: he foreshadowed the use of groups of symmetries, the topology of networks, decision theory and the theory of sets (he was, for instance, the first to draw Venn diagrams). Nearly alone among his contemporaries, he advocated for the beauty and importance of number theory. His work on prime numbers, in particular, set the stage for a golden age of mathematics that would follow decades later.

However, Euler's greatest legacy, in both pure and applied mathematics, was the field of analysis. Seventeenth-century mathematicians, culminating with Isaac Newton and his arch-enemy Gottfried Wilhelm Leibniz, had founded calculus — the study of the rates of change of quantities in time (differentials or derivatives) and the intimately related idea of areas between curves (integrals). Euler's analysis turned calculus into a powerful science and endowed mathematics and physics with their modern language and appearance.

The founders of calculus often grasped at concepts that they could not fully understand. The field relied on infinitesimals, which had a metaphysical aura so controversial that they were in part responsible for getting Galileo Galilei in hot water with the Catholic Church, according to historian of mathematics Amir Alexander (*Infinitesimal* (Oneworld, 2014); see *Nature* <http://doi.org/9hz>; 2014).

In Euler's time, that controversy was still far from resolved. There were no rigorous definitions of limits or of the continuum of the real numbers; neither was put onto solid foundations until the nineteenth century.

Just as he was unfazed by blindness, Euler did not let these troubles hinder his mathematical creativity. In his treatment of infinitesimals — used in differential and integral calculus and in adding up infinite series — he took an approach that Calinger describes as “happy-go-lucky”. Euler's pragmatism is reminiscent of the ‘shut up and calculate’ attitude of the vast majority of twentieth-century physicists towards quantum mechanics, setting problematic foundations aside to allow enormous progress in applications (D. Kaiser *Nature* **505**, 153–155; 2014). Euler's powers of intuition and his method of testing his hypotheses on special cases using his unparalleled calculation skills meant that his results



**Leonhard Euler: Mathematical Genius in the Enlightenment**  
RONALD S. CALINGER  
Princeton University Press: 2016.

RIGBY/SPL



➔ **NATURE.COM**  
Visit our blog on  
science in culture:  
[go.nature.com/2mn9yk](http://go.nature.com/2mn9yk)

were almost always right, says Calinger. At the centre of analysis, Euler placed the concept of differential equations: those that link a function and its derivatives, and in which the solution consists of calculating the function itself. (In celestial mechanics, for example, the functions can represent the trajectories of planets.) He came to be regarded as the “principal inventor” of the field, Calinger writes, and his work on analysis “displaced synthetic Euclidean geometry from its two-millennium primacy”.

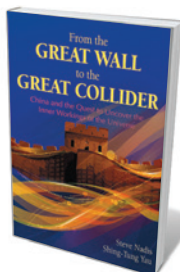
Euler demonstrated the power of this innovative science when he applied it to physical problems, such as the laws of the mechanics of solid bodies. In particular, he solved what many in the eighteenth century considered the most important open problem in astronomy: reconciling the complex motions of the Moon with Newton’s universal law of gravitation. This ‘three-body problem’ involves the interactions of the Sun, Moon and Earth, and is much harder than predicting one planet’s motion around the Sun. Some, including Euler, had suspected that Newton’s inverse-square law would break down in this crucial test, demanding the formulation of another theory. The problem had enormous practical importance: lunar motions could be used to calculate a vessel’s longitude at sea, and Euler was in the race to find a reliable method of doing so. (Eventually, precise timekeeping turned out to be a better solution.)

I have one quibble. The book’s strict chronological order means that it often reads as a sequence of disconnected summaries of Euler’s papers and correspondence, jumping from fundamental problems in algebra to ordering ink for his academy’s printing presses, often in the same paragraph. Still, fragmented as the narrative is, we manage to glimpse a personality. He was a man of integrity who — with few exceptions — gave credit where it was due and maintained a belief in “a harmony between written revelation and natural phenomena”. And although Calinger remarks on Euler’s perceived lack of “courtly manners”, we infer that this was really just a dearth of interest in flattering the nobility.

As a result, Euler never became head of the academies at which he worked, in Frederick the Great’s Berlin or Catherine the Great’s St Petersburg. No matter: his importance in the evolution of mathematics is clear. This impressively researched tome will be of great value to anyone with a serious interest in the history of mathematics and the Enlightenment. ■

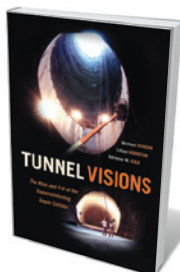
**Daide Castelvechi** is senior reporter on the physical sciences at Nature.

## Books in brief



### From the Great Wall to the Great Collider: China and the Quest to Uncover the Inner Workings of the Universe

Steve Nadis and Shing-Tung Yau INTERNATIONAL PRESS OF BOSTON (2015)  
The Large Hadron Collider (LHC) at Europe’s particle-physics lab, CERN, has witnessed game-changing discoveries, not least the Higgs boson in 2012. Now, rival ideas for successors are evolving (see *Nature* **511**, 394–395; 2014). In this forcefully argued history-cum-manifesto, physicist Shing-Tung Yau and writer Steve Nadis make the case for a “Great Collider” 100 kilometres in circumference to be built in China — an engineering marvel on a par with the Great Wall, but designed to lure hordes in for “rousing research collaboration”.



### Tunnel Visions: The Rise and Fall of the Superconducting Super Collider

Michael Riordan, Lillian Hoddeson and Adrienne W. Kolb UNIVERSITY OF CHICAGO PRESS (2015)

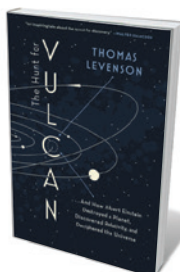
The termination of the Superconducting Super Collider project in 1993 sent more than US\$10 billion down the drain and left the US high-energy-physics community reeling. In this in-depth tome on that “epochal transition”, science historians Michael Riordan and Lillian Hoddeson, with Fermilab archivist Adrienne Kolb, cover all the bases leading to that bitter end — which, they conclude, was down to a “cold-war mindset” and the untenable cost of going it alone.



### Lady Byron and Her Daughters

Julia Markus W. W. NORTON (2015)

In the bicentenary of computer pioneer Ada Lovelace (R. Holmes *Nature* **525**, 30–32; 2015), it is salutary to remember her brilliant mother Annabella, estranged wife of volatile poet Lord Byron. Dubbed by him the “princess of parallelograms”, Annabella was a talented mathematician — but also a radical educational and social reformer, as Julia Markus reveals in this lucid biography. Annabella’s abolitionism sparked the admiration of Harriet Beecher Stowe, author of *Uncle Tom’s Cabin* (1852), who launched a fiery feminist defence of her in the face of virulent criticism from Byron’s hagiographers.



### The Hunt for Vulcan

Thomas Levenson RANDOM HOUSE (2015)

A science-fiction flavour clings to this real history of a nonexistent planet that sneaked into the annals of science, and the scientific icon who ushered it out again. Thomas Levenson wonderfully tells the story of Vulcan — the planet hypothesized (and ‘observed’) around 1860 to explain a wobble in Mercury’s orbit — as a frame for Albert Einstein’s general theory of relativity, which killed the putative planet stone dead. Looping through science history from Isaac Newton onwards, Levenson elegantly reveals the evolutionary nature of scientific thought, and the marvel of the revolution that Einstein wrought.



### We Are All Stardust

Stefan Klein, translated by Ross Benjamin EXPERIMENT (2015)

The dazzling clutch of scientific minds caught in mid-thought here makes for a read that provokes thought in its turn. Translated from German for the first time, this collection sees science writer Stefan Klein interview the likes of anthropologist Sarah Hrdy and astronomer Martin Rees. Delights abound. Rees uses the analogy of a department store to illustrate the emergence of life in a multiverse, while psychologist Alison Gopnik likens the intensity of babyhood to a first-time trip to Paris, revved up on Gauloises and espresso. **Barbara Kiser**





In *Evolution*, players use cards to develop species — and win points for their survival.

## EDUCATION

# How to win at evolution

Stuart West and helpers compare the cut and thrust of three games that explore life's greatest competition.

Evolution by natural selection is like a game: the winners are the organisms best at passing on their genes. Several board games are now available that exploit the strategies and uncertainties involved. We (two 12 year olds, a 16 year old, three graduate students, two postdocs and three professors) tested three: *Evolution* by North Star Games, *Evolution: Random Mutations* by Rightgames and *Terra Evolution: Tree of Life* by Mindwarrior Games. Our mission? To discover their potential for 'edutainment' — and as last-minute holiday gifts.

*Evolution* focuses on adaptation and inter-species interactions. You fashion your species, which competes for a food supply. Players are dealt cards allowing them to start new herbivore species, give an existing species a bigger population or bigger body, or add an advantage such as improved efficiency at foraging or the ability to store food as fat. Some cards turn a herbivore into a carnivore. Some help herbivores to avoid carnivores, with a shell or warning calls; other attributes, such as pack hunting and intelligence, help predators to evade these defences. Efficient herbivores get more food, but can be eaten by carnivores. Points are scored for eating, and for the species you have left at the end of the game. The player with the most points wins.

The same focus and similar mechanics

## Evolution

North Star Games: 2014.

## Terra Evolution: Tree of Life

Mindwarrior Games: 2015.

## Evolution: Random Mutations

Rightgames: 2013.

pervade *Evolution: Random Mutations*, but you are not allowed to know the ability on a card before you add it to a species. The game also has cards that derail a species' survival, such as developmental defects. Points accrue according to the number and complexity of the species you have at the end of the game, along with traits such as high body weight, which make a species need more food.

The more complex *Terra Evolution* offers a nice illustration of how the tree of life is built over time, and opens discussions about the repeatability of evolution. You start with seven continents on a planet and a simple, worm-like animal. You build a deck of cards allowing you to evolve complex organisms, fossilize others or attack rival players with catastrophes such as viruses or asteroids. Complexity in an animal can be a step towards greater complexity: fish lead to reptiles, and reptiles to birds or mammals. Or a complex animal might be an evolutionary dead-end, but give you a useful trait — wasps, for example, help you to attack other

players. The winner is the first to evolve both mammals and birds, and to score a certain number of fossil points.

We savoured all three games, but *Evolution* is our favourite by far. It looks amazing, with evocative artistry in everything from the cards to the little animal drumsticks that the predators eat. A gratuitous wooden dinosaur marks whose turn it is, and tactile, satiny bags store eaten food. The gameplay is simple to grasp, but can get very tactical. In particular, as with real evolution, the best strategy depends on what everyone else is doing. If there are a lot of herbivores, there is an advantage to being an efficient forager, with traits such as cooperation, but lots of herbivores also means a big advantage to becoming a carnivore. When carnivores appear, herbivores need defences, which carnivores try to get around — and so on, in a co-evolutionary dance.

*Evolution* features sophisticated biology. Traits can be put together in a dizzying array of combinations, so each game can be very different. The theme of evolution is not just tacked on: it drives play. A long neck gets you food that would not normally be available. Symbiosis means that you can only be eaten by a predator if it eats your host first. Horns do not stop you being eaten, but they damage predators, and so act as a deterrent.

*Evolution* captures key aspects of the evolutionary process and would work as a teaching aid for ages ten and up. It could also help older students to tackle specific topics, such as evolutionary arms races. *Random Mutations* is neither as attractive nor as fun, the welcome addition of parasites notwithstanding, but its small size could make it a good travel game. It is a little drier, although it does a great job of illustrating the random role of mutation, which can increase or decrease fitness.

*Terra Evolution* was the hardest to grasp — we had to play several times to get the hang of it. How you play early on can have long-term consequences, and the best strategy can vary with the stage of the game. We liked the ability to wreak cataclysms, although this worked best with fewer players. It would suit older players, who like more complex gameplay.

All the games, and especially *Evolution*, deftly capture how natural selection produces organisms adapted to their environments. I would love to see a board game exploring the dynamics of natural selection within species, through changes in gene frequency. But making population genetics fun is no mean feat. In the meantime, I am delighted to see so many attempts to drum up understanding of this greatest of all games. ■

Stuart West is Professor of Evolutionary Biology in the Department of Zoology at the University of Oxford, UK. He studies adaptation in organisms ranging from bacteria to humans.

e-mail: stuart.west@zoo.ox.ac.uk

# Correspondence

## Push to decarbonize cities after Paris talks

As the United Nations climate summit in Paris draws to a close, we suggest that cities should be a focus of action against climate change (see also D. G. Victor and J. P. Leape *Nature* **527**, 439–441; 2015). Covering just 3% of Earth's surface but housing more than half of its population, they account for 70% of global energy demands.

Cities are already ahead of nations on climate policy. Initiatives such as C40 Cities ([www.c40.org](http://www.c40.org)) and the World Mayors Council on Climate Change ([go.nature.com/yyqjre](http://go.nature.com/yyqjre)) help urban centres to integrate climate objectives into current policy and long-term planning. The Cities for Climate Protection campaign run by the global cities network ICLEI, for example, has prevented emissions equivalent to some 54 million tonnes of carbon dioxide from more than 1,000 cities (see [www.iclei.org](http://www.iclei.org)).

We shall still have to cope with factors that affect city infrastructures and supply chains, such as rising sea levels, accelerating migration from rural to urban areas, and more frequent and extreme weather events. Fundamental changes are needed in the way that we build and manage the urban environment (see [go.nature.com/q61pq5](http://go.nature.com/q61pq5)).

**Matthew Agarwala\*** *London School of Economics and Political Science, UK.*

[m.k.agarwala@lse.ac.uk](mailto:m.k.agarwala@lse.ac.uk)

*\*On behalf of 7 correspondents (see [go.nature.com/du4zld](http://go.nature.com/du4zld) for full list).*

## Continental targets for EU conservation

The European Union is currently evaluating the relevance of its environmental policies (see [go.nature.com/vkm9r7](http://go.nature.com/vkm9r7)). We have contributed to this by reviewing the 1,448 projects that were funded from 1992 to 2013 to the tune of €3 billion (US\$3.2 billion) by the EU

under LIFE Nature, its main financial tool for conservation programmes.

We find that a disproportionate amount of funding is directed towards species that are not under global threat and/or to regions of low conservation priority (unpublished data; see also [go.nature.com/omfvfn](http://go.nature.com/omfvfn)). In our view, the EU's conservation programme would benefit from more strategic planning and more flexibility when it comes to setting conservation priorities. Fixed lists, such as those setting conservation priorities in EU directives, can rapidly become outdated.

More funding is not necessarily the answer (see, by contrast, V. Kati *Conserv. Biol.* **29**, 260–270; 2015). We recommend that funds should be allocated according to continental and global needs, overriding national government interests. This would also help to prevent unnecessary duplication of efforts in different countries. In our view, continental conservation objectives, such as the EU's Biodiversity Strategy for 2020, need continental-scale plans.

**Virgilio Hermoso\*** *Catalonia Forest Sciences Centre (CTFC), Solsona, Spain.*

[virgilio.hermoso@gmail.com](mailto:virgilio.hermoso@gmail.com)

*\*On behalf of 4 correspondents (see [go.nature.com/ssgtgf](http://go.nature.com/ssgtgf) for full list).*

## Citizen scientists can aid diagnostics

Citizen scientists are an underrated source of observations on medical conditions. They frequently offer researchers a head start in the hunt for biomarkers (see, for example, the tentative identification of volatile indicators of early Parkinson's disease: [go.nature.com/wggoss](http://go.nature.com/wggoss)).

The precision and high-throughput capability of analytical technology drives most advances in clinical diagnostics (L. M. McShane *et al. Nature* **502**, 317–320;

2013). Analytical science and its subdiscipline metabolomics (the study of chemical fingerprints left by cellular processes) are also crucial for guiding clinical decisions (see [go.nature.com/l8pcde](http://go.nature.com/l8pcde)). These tools are set to be valuable for investigating and tapping into citizen scientists' previously unreported medical phenomena.

**Nicholas J. W. Rattray\***

*SYNBIOCHEM, University of Manchester, UK.*

[nicholas.rattray@manchester.ac.uk](mailto:nicholas.rattray@manchester.ac.uk)

*\*On behalf of 5 correspondents (see [go.nature.com/uy7w5n](http://go.nature.com/uy7w5n) for full list).*

## Europe must fund social sciences

The peer-reviewed Social Impact Open Repository (SIOR) was launched by the European Commission earlier this year to evaluate the social benefits of research ([www.ub.edu/sior](http://www.ub.edu/sior)). Its success is an argument against the commission's plans to eliminate the social sciences and humanities component from its Horizon 2020 research-funding programme.

Contributors of evidence of social impact to SIOR include the Informal Economy Monitoring Study of the WIEGO project (Women in Informal Employment: Globalizing and Organizing) based at Harvard University in Massachusetts. The project has helped women to organize labour, gain recognition and defend their rights in ten cities in Africa, Asia and Latin America (see [www.wiego.org](http://www.wiego.org)). Spain's Atapuerca project has enhanced economic growth and cultural development in the region where the first human bones were discovered (see E. Carbonell *et al. Nature* **452**, 465–469; 2008). And the ALACs project, which promotes citizenship through its Advocacy and Legal Advice Centres, has helped citizens to tackle corruption in the Balkan and Caucasus areas (see

[go.nature.com/qnmefy](http://go.nature.com/qnmefy)).

The open repository scores research on metrics such as social improvement, transferability to diverse populations or social contexts, and sustainability. SIOR data are also used to back research proposals.

**Ramon Flecha, Marta Soler-Gallart** *University of Barcelona, Spain.*

**Teresa Sordé** *University Autònoma de Barcelona, Spain.*  
[marta.soler@ub.edu](mailto:marta.soler@ub.edu)

## Science festivals preach to the choir

Science festivals are designed to expand the public's interest in science, but we find that this genre of science communication appeals mainly to a select clientele.

As part of our ongoing evaluation to determine how such activities might be improved (E. Jensen and N. Buckley *Public Underst. Sci.* **23**, 557–573; 2014 and E. Jensen *J. Sci. Commun.* **14**, Y05; 2015), we found that attendees at UK science festivals are more highly educated and economically advantaged than most of the population. Attendees are also disproportionately pro-science, and tend to be highly engaged in a wide range of cultural events.

This pattern reinforces existing sociocultural divides by benefiting privileged groups. Given their large investment in science communication, we recommend that scientists, universities and governments demand higher standards of inclusivity and use evidence-based practices at science festivals.

**Eric A. Jensen** *University of Warwick, Coventry, UK.*

**Eric B. Kennedy** *Arizona State University, Tempe, Arizona, USA.*

**Monae Verbeke** *Institute for Learning Innovation, Portland, Oregon, USA.*  
[ericbkennedy@asu.edu](mailto:ericbkennedy@asu.edu)



## CELL DIVISION

## A last-minute decision

Robust quantitative analyses of asymmetric division in certain cells in flies identify the major molecular players and, most interestingly, define a simple equation to explain this complex cellular process. [SEE LETTER P.280](#)

CAYETANO GONZALEZ

The development of multicellular organisms requires the generation of many cell types from a single precursor. In some cases, this is achieved by asymmetric cell divisions, in which the two daughters of a dividing cell receive different complements of intracellular 'fate' factors. Evolution has devised different mechanisms for asymmetric cell division<sup>1</sup>, some of which are rather counterintuitive — as in the sensory organ precursor (SOP) cells of the fruit fly. When SOPs divide, vesicles called Sara endosomes, which are loaded with cell-fate molecules, become restricted primarily to one daughter. But near the end of cell division, Sara endosomes are still found in the centre of SOPs, seemingly running the risk of being partitioned equally between the two daughter cells. It is only when cell division is close to completion that they find their way to one daughter<sup>2</sup>. On page 280 of this issue, Derivery *et al.*<sup>3</sup> explain how this last-minute decision works.

The authors recorded thousands of SOP divisions using cells in which molecules and structures of interest were engineered to fluoresce. Interpreting these data was a hugely daunting task, for several reasons. To begin with, the speed of events and the size of the

structures involved are not far from the temporal- and spatial-resolution limits of advanced microscopy. Furthermore, living cells are not all identical in size, shape and orientation. Signal-to-background ratios are often sub-optimal, and — let's face it — biology is often noisy, being subject to random fluctuations that obscure the real data. One video will never provide a hint of how this system works, and even 1,000 might prove no more informative.

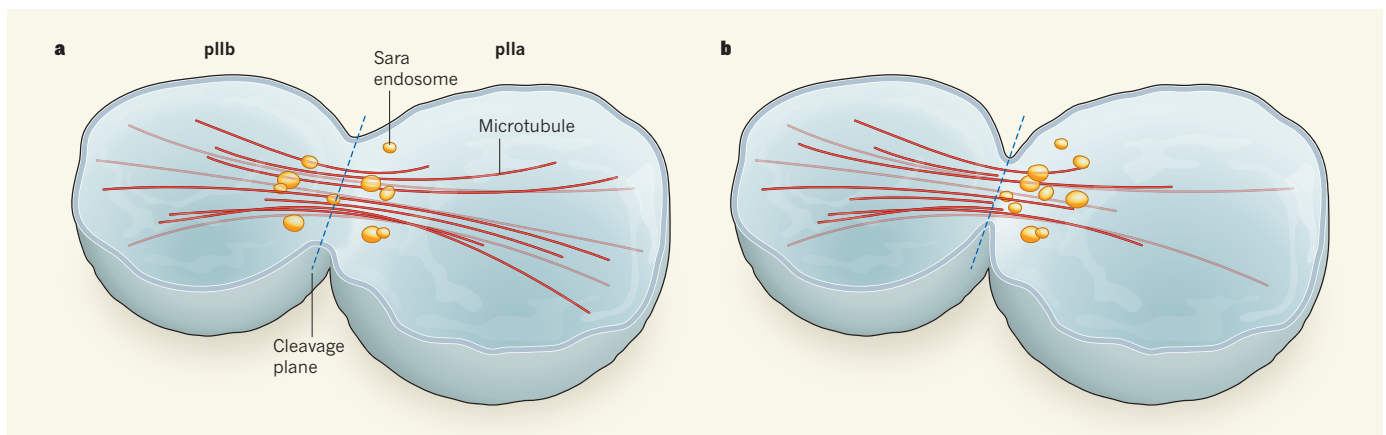
Derivery *et al.* found a creative way to circumvent these problems, making use of a contractile ring structure that cleaves the SOPs in two. Rather than trying to derive information from each individual video, the authors used the position of the ring and the onset of contraction as spatial and temporal reference points from which to align them, superimposing every recording to generate a movie of the 'average' SOP division (see [go.nature.com/ccnyjs](http://go.nature.com/ccnyjs)). This average cell is artificial, and does not exist in nature. Nonetheless, it is a feast of data, packed with precious, reproducible, spatio-temporal information about the cellular components involved in compartmentalizing Sara endosomes to one daughter.

Time-resolved density plots of different components of the average SOP revealed two distinct phases of division. A structure called the central spindle, which contains

many microtubules, is required for cell cleavage. During the first phase, the authors observed that both these microtubules and Sara endosomes, which move back and forth along the microtubules at a constant speed, are equally distributed between the two sides of the dividing cell (dubbed pIIa and pIIb). In the second phase, microtubule density is about 20% higher in the pIIb side, but Sara endosomes spend almost twice as much time in the pIIa side, and move beyond the end of microtubules more often in this direction. As a result, more than 80% of Sara endosomes end up in pIIa (Fig. 1).

Next, Derivery and colleagues demonstrated that the activity of just three proteins can explain these observations. Throughout the process, Sara endosomes are transported along the microtubules by the protein Klp98A — a relative of a mammalian motor protein that moves cargo to one end of microtubules<sup>4</sup>. The asymmetry of the microtubules at the central spindle is brought about by enrichment of the protein Patronin on the pIIb side. Patronin then protects microtubules that are oriented in a certain direction from the depolymerizing activity of the protein Klp10A (ref. 5). By contrast, Klp10A disassembles the unprotected microtubules on the pIIa side.

Remarkably, this complex intracellular choreography fits a rather simple equation — the



**Figure 1 | A mechanism for asymmetric cell division.** Derivery *et al.*<sup>3</sup> have defined the mechanism by which vesicles called Sara endosomes are apportioned asymmetrically into two daughter cells (pIIa and pIIb) during the division of sensory-organ precursor cells in the fruit fly. **a**, Towards the end of cell division, a spindle structure based on microtubules forms in the centre of

the cell. Sara endosomes move evenly up and down the microtubules, and the microtubules themselves are arrayed symmetrically about the cleavage plane, down which the cell will split. **b**, As division progresses, some microtubules are disassembled. Levels of disassembly are higher on the pIIa side of the cell. Sara endosomes spend more time in pIIa, accumulating on this side.

extent of endosome asymmetry is inversely proportional to microtubule asymmetry and exponentially amplified by several factors, including inverse microtubule asymmetry. This equation readily accounts for the counterintuitive observation that microtubule-binding endosomes accumulate on the side of the cell that has fewer microtubules, and explains how a small difference in microtubule density can translate into a large bias in endosome segregation. The authors confirmed that the equation works over a wide range of induced levels of central-spindle asymmetry, and defined each of its parameters experimentally.

Putting this equation to the test, Derivery and colleagues placed a Patronin-sequestering molecule on the pIIb side of a SOP, on a structural-protein network called the cortex that lines the interior of the membrane. This molecular trap draws Patronin from the pIIb side of the central spindle, making pIIb more susceptible to microtubule disassembly by Klp10A. The authors found that the microtubule network became denser in pIIa, and that most Sara endosomes were delivered to pIIb, as predicted by the equation. And, consistent with their hypothesis, placing the Patronin trap on the pIIa cortex affected neither microtubule asymmetry nor the location of Sara endosomes.

This study is filled with *in vitro* assays, mutant analyses, purpose-made computer code and more. Moreover, it is an excellent example of the power of using solidly quantified data to solve complex problems in molecular cell biology, applying equations when intuition will not do.

Given the ubiquitous nature of its components, the mechanism identified by Derivery and colleagues might not be restricted to Sara endosomes in dividing SOPs. It could be applicable to any process in which a molecular cargo is moved unidirectionally by a motor protein along microtubules that are arranged as an asymmetric bundle. Moreover, the evolutionary conservation of the proteins involved points to the possibility that this mechanism operates in other species. As the authors point out, one case that merits close examination is that of the branched projections of neurons that receive stimuli from other neurons. These projections contain many antiparallel microtubules, and vesicle transport plays a fundamental part in their function.

Future research will determine whether the authors' mechanism can be generalized to this, or to any other process, in mammalian cells. But, for now, we can say that fruit flies have once again delivered insight into a basic biological mechanism that may well be applicable to other organisms. ■

**Cayetano Gonzalez** is at the Institute for Research in Biomedicine (IRB Barcelona), the Barcelona Institute of Science and Technology, 08028 Barcelona, Spain; and at the Catalan Institution for Research and Advanced

Studies (ICREA), Barcelona.

e-mail: gonzalez@irbbarcelona.org

1. Knoblich, J. A. *Nature Rev. Mol. Cell Biol.* **2**, 11–20 (2001).
2. Coumilleau, F., Fürthauer, M., Knoblich, J. A. &

González-Gaitán, M. *Nature* **458**, 1051–1055 (2009).

3. Derivery, E. *et al.* *Nature* **528**, 280–285 (2015).
4. Hoepfner, S. *et al.* *Cell* **121**, 437–450 (2005).
5. Goodwin, S. S. & Vale, R. D. *Cell* **143**, 263–274 (2010).

## CATALYSIS

# The complexity of intimacy

**Catalysts that contain two types of active site split long hydrocarbon molecules into more-useful shorter ones. Research into controlling the nanoscale separation of the sites challenges accepted design rules for such catalysts. [SEE LETTER P.245](#)**

ROGER GLÄSER

**T**he petroleum industry produces jet fuel and diesel using a catalytic process called hydrocracking, in which long-chain hydrocarbons are broken into shorter, more-useful ones. More than 250 million tonnes of hydrocarbons are hydrocracked each year<sup>1</sup> (Fig. 1). The catalysts comprise an acidic microporous silicate called a zeolite and a noble metal, mixed together with a binder material<sup>1</sup>. On page 245 of this issue, Zečević *et al.*<sup>2</sup> report a major breakthrough in the design of hydrocracking catalysts, which greatly improves the selectivity of the process so that more of the desired products are formed. The results shatter the belief that the noble-metal sites and the acid sites must be as close together as possible for effective catalysis.

The distance between the acid sites and the noble-metal sites of hydrocracking catalysts must be below a maximum limit defined by the

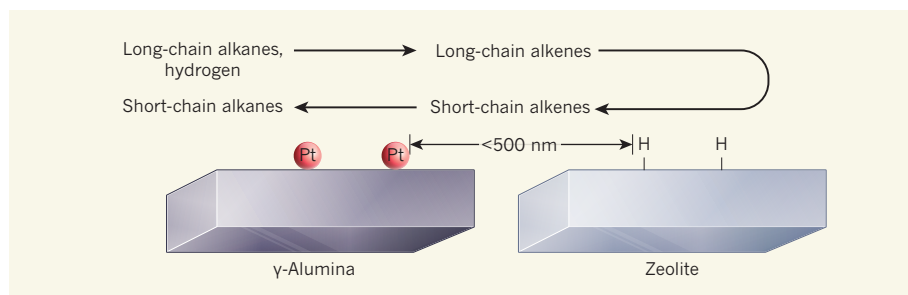
'intimacy criterion', which was first reported<sup>3</sup> in 1962. The metal sites catalyse dehydrogenation reactions that convert the alkane reactants into unsaturated hydrocarbons called alkenes, which must then diffuse to an acid site to undergo isomerization and cracking (cleavage into shorter molecules; Fig. 2). If the metal sites are too far from the acid sites, then catalytic activity decreases — satisfying the intimacy criterion therefore overcomes diffusive limitations of the catalytic process.

Previous studies of the intimacy criterion have examined inter-site distances only at or above the micrometre scale. Zečević *et al.* now report two catalysts in which platinum was controllably placed at nanometre-scale distances from the acid sites. The catalysts contain platinum particles of around 3 nanometres in diameter, with a narrow particle-size distribution, so that the metal was selectively supported on either the zeolite or the binder (γ-alumina); the overall platinum content in



**Figure 1** | A hydrocracking facility in Gdansk, Poland.





**Figure 2 | Spacing of active sites in hydrocracking catalysts.** Alkane molecules that have long carbon chains are reacted with hydrogen in the presence of a solid catalyst to produce shorter alkane molecules, a process called hydrocracking. The catalyst consists of a noble metal (such as platinum, Pt) on a  $\gamma$ -alumina support, and a porous solid called a zeolite, which contains acid sites (indicated by hydrogen atoms). The metal converts long-chain alkanes into long-chain alkene intermediates, which then diffuse to the acid sites, where they are isomerized and cracked into short-chain alkenes. These alkenes diffuse back to the platinum sites, where they react with hydrogen to form short-chain alkanes. Alkene diffusion between the metal and the acid sites is rate limiting if the distance between the sites is larger than 500 nanometres. Zečević *et al.*<sup>2</sup> have prepared catalysts with improved catalytic selectivity compared with conventional catalysts, by controlling the distance of the acid and metal sites on the nanoscale.

the two catalysts was comparable, and less than 1% by weight. Key to this nanoscale control is the use of specific interactions between the noble metal's precursor and its support. One of the authors' major achievements was to use advanced electron-microscopy techniques to prove that the composite catalysts had identical structural parameters, except for the location of the platinum sites.

Another highlight is the design of the catalytic experiments. By using hydrocarbon feeds with molecules of different carbon-chain length and bulkiness, the authors were able to distinguish the influence of site proximity on catalytic activity and selectivity. They show that conversion of *n*-decane and the longer *n*-nonadecane into cracking products is the same when using the catalyst in which  $\gamma$ -alumina supports the platinum — that is, the intimacy criterion for activity is fulfilled for these reactions. But conversion of pristane, which has bulky molecules, is lower than for the other feeds, which indicates that diffusion of reaction intermediates between the active metal and acid sites limits catalytic activity in this case.

Strikingly, however, more of the desired isomerization products and fewer unwanted side products of cracking are formed from nonadecane than from decane when using the platinum-on-alumina catalyst than when using the platinum-on-zeolite catalyst — that is, the selectivity of the reaction for desirable products is higher when using the former catalyst. But the metal and acid sites are closer together in the latter catalyst than in the former, and so the difference in product selectivity apparently contradicts conventional understanding of the intimacy criterion. This can be explained if longer hydrocarbon molecules spend a greater amount of time in the zeolite's micropores in the second catalyst, and therefore undergo multiple cracking on the acid sites.

When the platinum particles reside on  $\gamma$ -alumina, still in the vicinity of acid sites, the

intermediate alkenes formed during hydrocracking diffuse to the zeolite and rapidly isomerize. The authors propose that reactions of the alkenes occur close to the outer surface of the zeolite crystals, from where they can easily be desorbed and transported back to the metal sites on the  $\gamma$ -alumina. This mechanism was previously suggested<sup>4</sup> by these authors to be responsible for the reactions of long-chain hydrocarbons on the outer layers of zeolites that have medium-sized pores, and is now convincingly shown to be valid for the large-pore zeolite studied by Zečević and co-workers.

It is therefore evident that catalytic activity and selectivity depend not only on the distance between the active sites and on the molecular dimensions of the reactants, but also on the accessibility of the sites to molecules and on the transfer rate of molecules between sites. These factors might be even more important when more-complex feedstocks, such as fats and oils from renewable resources, are processed, or when catalysts contain more than two kinds of active site, or complicated pore architectures<sup>5</sup>.

HIV

## Cure by killing

**Two bi-specific protein constructs have been designed that direct the body's T cells to kill HIV-infected cells. The feat provides a step on the path to removing the latent virus reservoir that persists in patients on antiretroviral therapy.**

DOUGLAS D. RICHMAN

The development of combination antiretroviral therapy to suppress HIV infection and its complications has been a major achievement of modern medicine. However, these drugs do not eradicate the virus; they only suppress its productive replication

In such cases, proving the benefits of nanoscale site intimacy will be considerably more challenging than for the relatively simple hydrocracking of single-component hydrocarbon feeds over a bifunctional platinum-zeolite catalyst studied here.

Two major lessons can be learnt from the present study. First, preparation strategies, modern visualization techniques and in-depth catalytic studies must be combined and directly correlated to improve the efficiency of complex multifunctional catalysts. Second, mass-transfer effects are often prominent and may even govern the conversion of reactants on solid catalysts<sup>6</sup>. They should therefore be carefully considered and optimized — not only for the bulk catalyst or the entire reaction chamber, but also for catalytic sites at the nanoscale. This presents a challenge for both experimentalists and theoreticians. More broadly, Zečević and co-workers' method for controlling the nanoscale structure of hydrocracking catalysts may benefit several other processes that use solid catalysts, including the conversion of renewable resources (such as fats, oils or biomass) into more-valuable products, or 'upgrading' heavy hydrocarbons to more-useful compounds. ■

**Roger Gläser** is at the Institute of Chemical Technology, Universität Leipzig, Leipzig 04103, Germany.  
e-mail: roger.glaeser@uni-leipzig.de

1. van Veen, J. A. R., Minderhoud, J. K., Huve, L. G. & Stork, W. H. J. in *Handbook of Heterogeneous Catalysis* Vol. 6 (eds Ertl, G., Knözinger, H., Weitkamp, J. & Schüth, F.) 2778–2808 (Wiley-VCH, 2008).
2. Zečević, J. *et al. Nature* **528**, 245–248 (2015).
3. Weisz, P. B. *Adv. Catal.* **13**, 137–190 (1962).
4. Martens, J. A. *et al. Angew. Chem. Int. Edn* **34**, 2528–2530 (1995).
5. Garcia-Martinez, J. & Li, K. (eds) *Mesoporous Zeolites* (Wiley-VCH, 2015).
6. Kärger, J., Ruthven, D. M. & Theodorou, D. N. *Diffusion in Nanoporous Materials* Vol. 2, 807–837 (Wiley-VCH, 2012).

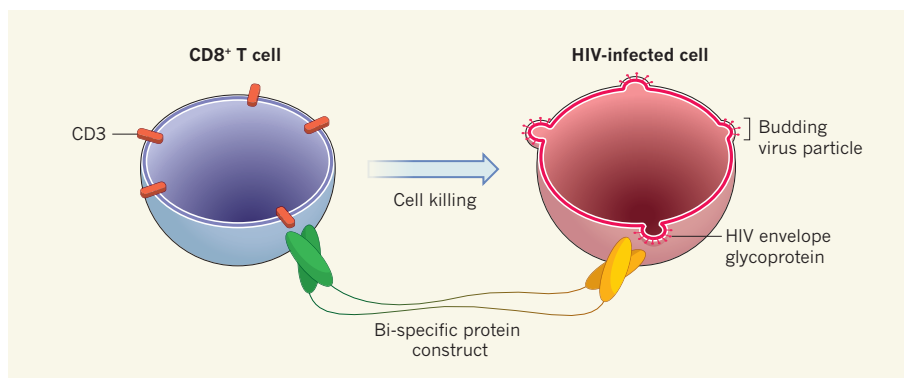
cycle in infected cells. This cycle involves the integration of HIV DNA into the genome of a host cell, and the generation of new viral particles that are released by budding from the cell. A small proportion of cells survives the cell-lytic consequences of infection, and goes on to form a latent reservoir of cells that have HIV DNA integrated into host-cell chromosomes.

This reservoir can rekindle a raging infection if antiretroviral therapy is interrupted. Only strategies that eradicate the reservoir will achieve a full cure and allow patients to discontinue a treatment that is costly, inconvenient and has side effects. Two publications, from Sung *et al.*<sup>1</sup> in *The Journal of Clinical Investigation* and Pegu *et al.*<sup>2</sup> in *Nature Communications*, describe innovative molecular constructs designed to selectively kill these rare, latently infected cells.

To be targeted by drugs, infected cells need to display evidence that they are harbouring the virus. For latently infected cells, this requires the virus to be activated to restart its replication cycle. Over the past few years, approaches to the selective activation of latently infected cells have used drugs such as histone deacetylase inhibitors, which are designed to induce the transcription of viral RNA. However, only a small minority of cells responds to these compounds<sup>3–5</sup>. Moreover, destruction of the activated cells may rely on the patient's immune responses, but these are already impaired and could be targeted to act against the initially infecting virus, which may have mutated to escape immune recognition<sup>6</sup>.

In attempts to circumvent these problems, Sung *et al.* and Pegu *et al.* describe conceptually similar but different protein constructs that combine the binding specificities of two different antibodies (Fig. 1). One of the antibodies binds to a broad spectrum of HIV envelope proteins, which are displayed on the surface of actively infected cells, typically CD4<sup>+</sup> T cells of the immune system; the other binds to the molecule CD3, which is displayed on the surface of all T cells. The rationale is to direct the CD8<sup>+</sup> subset of T cells, which has cell-killing (cytotoxic) ability, to kill latently infected cells that have been induced to express envelope proteins. The engagement of CD3 means that any CD8<sup>+</sup> T cell can be targeted to the infected cells, obviating the need for the T cell to specifically bind to HIV surface glycoproteins (antigens).

Both groups show that, *in vitro*, their constructs work as intended, inducing direct killing of cells that express envelope proteins, independently of the CD8<sup>+</sup> T cells' antigen specificity. The DART (dual-affinity re-targeting) construct described by Sung *et al.* also worked *ex vivo*: it killed cells taken from patients whose infection was well suppressed with antiretroviral therapy after the cells had been induced to express envelope proteins (through exposure to the protein phytohaemagglutinin or the histone deacetylase inhibitor vorinostat). Pegu *et al.* show that the CD3-binding arm of their construct not only activates cytotoxic CD8<sup>+</sup> T cells, but also serves to activate latently infected CD4<sup>+</sup> T cells to express HIV envelope proteins, thus permitting them to be killed without other inducing factors. However, additional studies will be needed to provide evidence for substantial killing of patient cells *ex vivo* using this construct.



**Figure 1 | Bi-specific constructs.** Sung *et al.*<sup>1</sup> and Pegu *et al.*<sup>2</sup> present two different protein constructs that combine sections of two monoclonal antibodies: one antibody binds to the CD3 molecule expressed on the surface of all T cells of the immune system, and the other binds to envelope proteins of the HIV virus, which are expressed on the surface of infected cells. These constructs direct CD8<sup>+</sup> (cytotoxic) T cells to kill HIV-infected cells, regardless of whether receptors on the CD8<sup>+</sup> T cell have specificity for the virus. This targeting is designed to increase the efficiency of the anti-HIV immune response and to help clear rare, latently infected cells.

Many previous investigations have generated bi-specific antibodies against various targets, including antigens from HIV and from tumour cells. Bi-specific antibodies for killing HIV-infected cells were first described in 1991, before potent antiretroviral drug combinations existed and before a cure for HIV was a conceivable goal<sup>7,8</sup>, and bi-specific constructs have been shown to kill cancer cells in non-Hodgkin's lymphoma<sup>9</sup>. However, the translation of such antibodies from an innovative molecular construct with *in vitro* activity to an effective treatment is a daunting process — and translation of the new constructs will be no exception.

There are several biological features that represent hurdles to any strategy for HIV cure<sup>10</sup>. Only one in a million CD4<sup>+</sup> T cells in HIV-infected individuals is latently infected, which means that these cells are a rare target and are difficult to detect and measure. So far, no agents have been identified that induce HIV antigen expression in the majority of latently infected cells while not affecting non-infected cells. Furthermore, the latent reservoir may include cells in anatomical compartments that large protein constructs cannot penetrate, such as the central nervous system and genital tract.

Moreover, every promising candidate compound faces substantial developmental hurdles before it can become an effective drug. First, the efficacy seen in *in vitro* models may not be replicated *in vivo*. Drugs can be kept at a constant concentration in cell culture, but this is difficult to achieve *in vivo*, where absorption, distribution and clearance of the compound occurs. Second, the target cells and cytotoxic cells may also not be as dense or uniformly distributed in tissues as in cell culture. A third consideration is that no envelope-binding antibody will bind to all HIV envelopes, and the possibility of the virus mutating to escape recognition by the construct must be considered.

Off-target effects may also be a concern.

Pegu and colleagues' bi-specific antibody induced a rapid drop in the number of CD3-expressing cells in the bloodstream when infused into infected rhesus macaques, probably owing to the activation and redistribution of the cells. The authors' construct also activates uninfected CD4<sup>+</sup> T cells. Finally, protein constructs can induce immune responses, as documented by Pegu and colleagues in monkeys. The anti-construct antibodies generated during such responses might be toxic, and will certainly result in diminishing activity of the drug.

Research into a cure for HIV is in the early stages of a long and difficult path, and all innovative options should be welcomed and investigated. Monoclonal antibodies have been remarkably successful in treating many medical conditions, and constructs that exploit their specificity to generate bi-functional capabilities — such as those presented by Sung *et al.* and Pegu *et al.* — are one such innovation worthy of further pursuit. ■

**Douglas D. Richman** is at the Center for AIDS Research, University of California San Diego, La Jolla, California 92093-0679, USA, and at the Veteran Affairs San Diego Healthcare System, San Diego.  
e-mail: drichman@ucsd.edu

1. Sung, J. A. M. *et al.* *J. Clin. Invest.* **125**, 4077–4090 (2015).
2. Pegu, A. *et al.* *Nature Commun.* **6**, 8447 (2015).
3. Archin, N. M. *et al.* *Nature* **487**, 482–485 (2012).
4. Elliott, J. H. *et al.* *PLoS Pathog.* **10**, e1004473 (2014).
5. Rasmussen, T. A. *et al.* *Lancet HIV* **1**, e13–e21 (2014).
6. Deng, K. *et al.* *Nature* **517**, 381–385 (2015).
7. Berg, J. *et al.* *Proc. Natl Acad. Sci. USA* **88**, 4723–4727 (1991).
8. Trautnecker, A., Lanzavecchia, A. & Karjalainen, K. *EMBO J.* **10**, 3655–3659 (1991).
9. Bargou, R. *et al.* *Science* **321**, 974–977 (2008).
10. Richman, D. D. *et al.* *Science* **323**, 1304–1307 (2009).

This article was published online on 2 December 2015.



## CLIMATE SCIENCE

# The Sun and the rain

**A study shows that, as Earth warms, global precipitation will increase by less than many models predict, because of increases in the amount of near-infrared sunlight absorbed by water vapour. SEE LETTER P.249**

STEVEN SHERWOOD

Earth's hydrologic cycle is expected to intensify as the planet warms, bringing increased average rainfall and evaporation. But calculations made by different climate models do not agree on how strong this effect will be, predicting anything from 10% to 16% more precipitation for 5 °C of global warming. On page 249 of this issue, DeAngelis *et al.*<sup>1</sup> report the reason for a large part of this disagreement: varying approximations in the models affect how strongly the absorption of sunlight by the atmosphere depends on the amount of water vapour. Happily, the authors use observed covariations of water vapour and radiation to work out the correct dependence, narrowing the spread of global precipitation projections — as it happens, towards the low end of the estimates.

Air is not as transparent to sunlight as it seems. Visible light mostly passes through clear air without being absorbed (Fig. 1), but accounts for only about half of the total energy in the solar spectrum. Much of the rest is near-infrared light, which can be absorbed by water vapour. (For liquid water, this absorption spreads into the red part of the visible spectrum, explaining the wonderful blue colour of a tropical lagoon or a swimming pool.) So not all sunlight absorbed by Earth heats the

ground; some of it heats the air as it is absorbed by water vapour, pollutants or ozone<sup>2</sup>.

Sunlight is absorbed mostly at Earth's surface, and this absorption drives the hydrologic cycle. By contrast, the infrared emission that returns this energy back to space originates mostly in the atmosphere. The atmosphere must close the circuit, and does so mainly by transporting surface-produced water vapour aloft, where it condenses, transferring latent heat upwards<sup>3</sup>. Any substance that absorbs sunlight in the atmosphere will partly short-circuit this process, weakening the hydrologic cycle, as will anything that increases the fraction of infrared emission to space that came from the surface.

This fraction is reduced by water vapour and carbon dioxide because of their greenhouse effect, so these gases strengthen the hydrologic cycle (for a given rate of thermal emission to space). Moreover, a warm atmosphere contains more water vapour than a cold one — about 7% more for each 1 °C rise (ref. 4). DeAngelis *et al.* report that this effect alone would cause more than a 15% increase in precipitation in a world carbon-warmed by 5 °C. However, the extra water vapour absorbs sunlight, which pushes the cycle in the opposite direction. The magnitude of this opposing push varies substantially between models<sup>5</sup>. DeAngelis *et al.* show that its variation is almost entirely due to differences

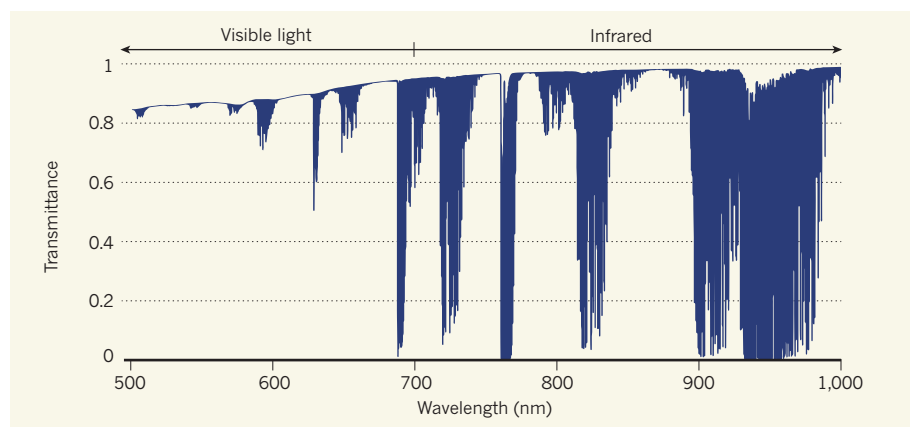
in the sensitivity of the model-calculated absorption of sunlight to a given increase in water vapour.

Calculating energy transfers by radiation through an atmosphere that has richly varying spectral properties (Fig. 1) is one of the most computationally expensive tasks for a climate model, so modellers design schemes that divide the spectrum into manageable chunks<sup>6</sup>. The primary absorption bands are probably well represented, but thousands of weak absorption lines throughout the spectrum may be overlooked or oversimplified, and there is much still to learn about such lines. These lines can collectively become a substantial player at high concentrations of absorbing gases<sup>7</sup>. This is a concern because modellers now routinely consider exotic palaeoclimates<sup>8</sup> or scenarios in which levels of atmospheric CO<sub>2</sub> have increased to four times<sup>6</sup> (or more<sup>7</sup>) the pre-industrial level, and for these the radiation schemes used in models may not be up to scratch. Now we know that such problems also greatly affect simulations of the hydrologic cycle<sup>1,5,6,9</sup>.

The good news is that this problem can be fixed. DeAngelis *et al.* look at what type of radiation scheme each model uses, and identify those that are not performing well. Not surprisingly, the worst offenders are old schemes that have coarse treatments based on older spectroscopic databases. Modellers will need to seriously consider upgrading these.

It is remarkable for a paper not only to identify a useful link between observable behaviour in today's climate and a crucial aspect of global climate change, providing a physical explanation, but also to trace that link back to a specific scheme within models. Such links, sometimes called emergent constraints, are now a hot topic in efforts to narrow the known uncertainty (model spread) in predictions of global warming. They should take centre stage in any efforts to 'weight'<sup>10</sup> the predictions of some models over others. But most emergent constraints reported so far either lack a clear physical explanation or fail to significantly narrow the uncertainty, either because the relationship is insufficiently strong or because there are not enough relevant observations to exploit it<sup>11</sup>. DeAngelis and colleagues provide an example of what such efforts should aspire to.

Their result is impressive, but its value for our understanding of climate change is more theoretical than practical. The main impacts from global warming will depend on regional changes in the amount and intermittency of precipitation, rather than on the global, time-averaged amount. Nonetheless, the finding is a wake-up call that radiative transfers of energy are a problem that is not completely solved, and that the partitioning of radiation flows in the atmosphere is as important for precipitation as the total heating of the Earth system is for its temperature. ■



**Figure 1 | Transmittance of sunlight by Earth's atmosphere.** The visible part of the solar light spectrum is mostly transmitted through the atmosphere without being absorbed (lines indicate parts of the spectrum that are absorbed; zero transmittance is equivalent to 100% absorption or scattering). However, the rest of the spectrum (mostly in the infrared) is absorbed much more strongly. Different climate models use different approximations of this spectral structure. DeAngelis *et al.*<sup>1</sup> report that these differences affect how the absorption of sunlight by the atmosphere varies with the amount of atmospheric water vapour in the models, and that this accounts for the large variation in predictions of how much average worldwide precipitation will increase with global warming. Adapted from [www.pole-ether.fr/tapas](http://www.pole-ether.fr/tapas)

Steven Sherwood is at the Climate Change Research Centre, and the ARC Centre for Climate System Science, University of New South Wales, Sydney, New South Wales 2052, Australia.  
e-mail: s.sherwood@unsw.edu.au

1. DeAngelis, A. M., Qu, X., Zelinka, M. D. & Hall, A.

*Nature* **528**, 249–253 (2015).  
2. Horvath, H. *Atmos. Environ. A* **27**, 293–317 (1993).  
3. O’Gorman, P. A., Allan, R. P., Byrne, M. P. & Previdi, M. *Surv. Geophys.* **33**, 585–608 (2012).  
4. Sherwood, S. C. *et al.* *J. Geophys. Res.* **115**, D09104 (2010).  
5. Takahashi, K. *J. Clim.* **22**, 5667–5675 (2009).  
6. Pincus, R. *et al.* *Geophys. Res. Lett.* **42**, 5485–5492 (2015).

7. Goldblatt, C., Lenton, T. M. & Watson, A. J. *Q. J. R. Meteorol. Soc.* **135**, 619–633 (2009).  
8. Halevy, I., Pierrehumbert, R. T. & Schrag, D. P. *J. Geophys. Res.* **114**, D18112 (2009).  
9. Fieldier, B. & Collins, W. D. *Geophys. Res. Lett.* **42**, 8749–8757 (2015).  
10. Klocke, D., Pincus, R. & Quaas, J. *J. Clim.* **24**, 6092–6099 (2011).  
11. Klein, S. A. & Hall, A. *Curr. Clim. Change Rep.* **1**, 276–287 (2015).

## CELL BIOLOGY

# Architecture of a protein entry gate

**The TOM complex guides precursor proteins from the cell’s cytosolic fluid into organelles called mitochondria. Biochemical analyses reveal the architecture of this complex and show how precursor proteins pass through its narrow pores.**

DEJANA MOKRANJAC & WALTER NEUPERT

Intracellular organelles called mitochondria are bounded by an inner and an outer membrane and house their own genome — two reminders that they arose from a bacterium that was engulfed by a host cell. As mitochondria evolved, most of their genome was transferred to the nucleus, and mitochondrial precursor proteins are now imported into mitochondria from the cell’s cytosolic fluid. Writing in *Science*, Shiota *et al.*<sup>1</sup> describe the molecular architecture of the main entry gate for these proteins, the TOM complex, in its native environment, achieving a milestone in the study of mitochondrial-import pathways.

The pathways that import mitochondrial precursor proteins are amazingly intricate<sup>2</sup>, involving some 40 proteins organized into 10 complexes called mitochondrial protein translocases. Together, the translocases ensure that more than 1,000 proteins, of diverse types, are correctly recognized, imported and sorted to the right place inside mitochondria. Much progress has been made in identifying the components of each translocase<sup>2</sup>, but atomic-resolution structures, which are needed for a deeper mechanistic understanding, are available for only a few soluble complexes and, with some translocases, for individual protein domains. The structural characterization of intact translocases is far from trivial, because they are complex, of low abundance, unstable and contain many membrane-spanning components.

The TOM complex recognizes

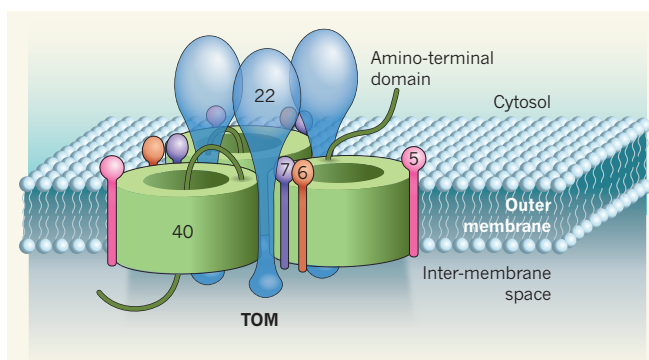
and guides mitochondrial proteins across the organelle’s outer membrane. Many experiments<sup>2</sup> have demonstrated that the complex forms pores in the outer membrane that are each made up of a channel-forming  $\beta$ -barrel protein called Tom40 and six proteins containing single  $\alpha$ -helical membrane-spanning segments — three receptors (Tom22, Tom20 and Tom70) that recognize mitochondrial precursor proteins for import and three subunits (Tom5, Tom6 and Tom7) that are functionally less well defined. Structural analyses have suggested that the TOM complex has two or three pores<sup>3,4</sup>. But although these studies made it relatively clear that Tom40 is the major pore-forming component, they were not conducted

at high-enough resolution to identify the positions of the other subunits, nor to define whether precursor proteins traverse the membrane through the Tom40  $\beta$ -barrels or through the spaces formed between the  $\beta$ -barrel pores.

Shiota *et al.* incorporated an unnatural amino acid into various sites in Tom40 *in vivo*. On activation by light<sup>5</sup>, this amino acid forms crosslinks with neighbouring proteins, revealing the positions of each subunit in the complex relative to Tom40. The authors overlaid these crosslinking data on a computational model of the structure of the Tom40  $\beta$ -barrel, which was generated using previously obtained data<sup>6–8</sup>. On the basis of the observed crosslinks, the researchers proposed a three-pore model for the TOM complex, in which three Tom40  $\beta$ -barrels (each consisting of 19  $\beta$ -strands that span the outer membrane) are bridged by three Tom22 molecules, such that each Tom22 binds to two  $\beta$ -barrels. Tom5, Tom6 and Tom7 are positioned at the outer face of the  $\beta$ -barrel (Fig. 1). Importantly, Shiota and colleagues also analysed interactions between Tom40 and examples of two major types of precursor protein. They observed that, during import, only amino-acid residues facing the interior of the  $\beta$ -barrel are crosslinked to the precursors. These data strongly support the idea that precursor proteins cross the outer membrane through  $\beta$ -barrels.

The two types of precursor analysed — one containing positively charged targeting signals, one hydrophobic — seemed to take different routes through Tom40. Whether all precursor proteins are imported through the  $\beta$ -barrel, or whether some take alternative routes, remains to be clarified. It will be interesting to investigate the import routes of proteins that were previously suggested<sup>2</sup> not to use Tom40. It will also now be possible to analyse the folding state of precursors as they pass through the narrow pore of the  $\beta$ -barrel.

The functional significance of the three pores in the complex remains unclear. Previous biochemical work<sup>9,10</sup> indicates that the TOM complex undergoes structural rearrangement during protein import. It could be that all the pores are identical and simultaneously active, or that two



**Figure 1 | Architecture of the TOM complex.** The TOM protein complex transports precursor proteins from their site of synthesis in the cell’s cytosolic fluid across the outer membrane of organelles called mitochondria. Shiota *et al.*<sup>1</sup> defined relative positions of subunits in the complex. The complex is made of three Tom40 subunits, which are connected to one another by Tom22 subunits. Each Tom40 subunit forms a pore across the membrane through which precursor proteins can pass from the cytosol into the space between the outer and inner mitochondrial membranes, possibly aided by the movement of Tom40’s amino-terminal domain. Tom5, Tom6 and Tom7 subunits are located around the periphery of the pores. The relative positions of the receptor proteins Tom20 and Tom70, which bind the precursor proteins, could not be precisely determined in this study and so are not included in this illustration.



inactive pores maintain the conformation of one active pore. It could also be that several pores are needed to attract the protein-synthesis machinery, thus ensuring that precursors are swiftly imported after their synthesis. It will be interesting to see whether molecular modification<sup>11</sup>, such as phosphorylation, of the TOM complex influences its multimeric state or its conformation.

Shiota and colleagues' model provides clues to a possible molecular mechanism underlying the structural rearrangement of the TOM complex. Tom22 contacts Tom40 where the first and last  $\beta$ -strand meet, an optimal site for Tom22 to regulate the conformation of the  $\beta$ -barrel. Notably, the first and last  $\beta$ -strands of Tom40 are parallel to one another. Lateral opening (in which these  $\beta$ -strands part, opening the barrel lengthways across the membrane) has been reported for  $\beta$ -barrels sealed by anti-parallel strands<sup>12</sup>. Because the stability of parallel  $\beta$ -strands is lower than that of antiparallel strands, the conformational flexibility of Tom40 and thereby of the entire complex might be higher than previously assumed. This could reflect the role of the complex in transporting diverse types of precursor protein.

Finally, the authors' model suggests that the amino-terminal segment of Tom40 can traverse the  $\beta$ -barrel from the cytosolic side to the inter-membrane space. Perhaps this segment blocks the pore, retracting before or during protein import. Alternatively, it might have an active role in moving precursor proteins across the barrel. The loops connecting individual  $\beta$ -strands are also likely to be involved in protein movement, as in related bacterial transporters<sup>13</sup>.

Future experiments should address how the interplay between precursor proteins and TOM-complex receptors influences the architecture of the entire complex. Does it undergo conformational changes as precursors are moved? Evidence suggests that some precursor proteins can also exit mitochondria through the TOM complex<sup>14</sup>, raising questions about how directionality is achieved. Because different types of precursor protein preferentially bind to different receptors, it is likely that the TOM complex responds to each type in a different manner.

The TOM complex cooperates with different downstream translocases to ensure that precursor proteins reach the correct location. Perhaps binding to downstream translocases changes the conformation of the TOM complex to optimize the efficiency with which specific types of precursor protein are imported, even before they bind to the complex<sup>15</sup>. The new model will allow a directed analysis of the interplay between the TOM complex and downstream translocases, building on previous studies<sup>1,10,15</sup>.

Shiota and colleagues' pioneering model will certainly inspire many hypothesis-driven analyses. Moreover, it will provide impetus for further structural and functional studies

of the molecular mechanisms that underlie mitochondrial-protein import. ■

**Dejana Mokranjac** is in the Department of Physiological Chemistry, Biomedical Center Munich, Ludwig-Maximilians University, 82152 Martinsried, Germany. **Walter Neupert** is at the Max Planck Institute of Biochemistry, 82152 Martinsried, Germany. e-mails: [dejana.mokranjac@med.lmu.de](mailto:dejana.mokranjac@med.lmu.de); [neupert@biochem.mpg.de](mailto:neupert@biochem.mpg.de)

1. Shiota, T. *et al. Science* **349**, 1544–1548 (2015).
2. Neupert, W. J. *Mol. Biol.* **427**, 1135–1158 (2015).
3. Künkele, K. P. *et al. Cell* **93**, 1009–1019 (1998).
4. Model, K., Meisinger, C. & Kühlbrandt, W. J. *Mol. Biol.* **383**, 1049–1057 (2008).
5. Liu, C. C. & Schultz, P. G. *Annu. Rev. Biochem.* **79**, 413–444 (2010).

6. Ujwal, R. *et al. Proc. Natl Acad. Sci. USA* **105**, 17742–17747 (2008).
7. Qiu, J. *et al. Cell* **154**, 596–608 (2013).
8. Lackey, S. W. K. *et al. J. Biol. Chem.* **289**, 21640–21650 (2014).
9. Rapaport, D. *et al. Mol. Cell. Biol.* **18**, 5256–5262 (1998).
10. Shiota, T., Mabuchi, H., Tanaka-Yamano, S., Yamano, K. & Endo, T. *Proc. Natl Acad. Sci. USA* **108**, 15179–15183 (2011).
11. Schmidt, O. *et al. Cell* **144**, 227–239 (2011).
12. Noinaj, N., Kusak, A. J., Balusek, C., Gumbart, J. C. & Buchanan, S. K. *Structure* **22**, 1055–1062 (2014).
13. Baud, C. *et al. Nature Commun.* **5**, 5271 (2014).
14. Bragoszewski, P. *et al. Proc. Natl Acad. Sci. USA* **112**, 7713–7718 (2015).
15. Waagemann, K., Popov-Čeleketić, D., Neupert, W., Azem, A. & Mokranjac, D. J. *Mol. Biol.* **427**, 1075–1084 (2015).

This article was published online on 25 November 2015.

## PLANETARY SCIENCE

# How the Solar System didn't form

**Standard planet-formation models have been unable to reconstruct the distributions of the Solar System's small, rocky planets and asteroids in the same simulation. A new analysis suggests that it cannot be done.**

KLEOMENIS TSIGANIS

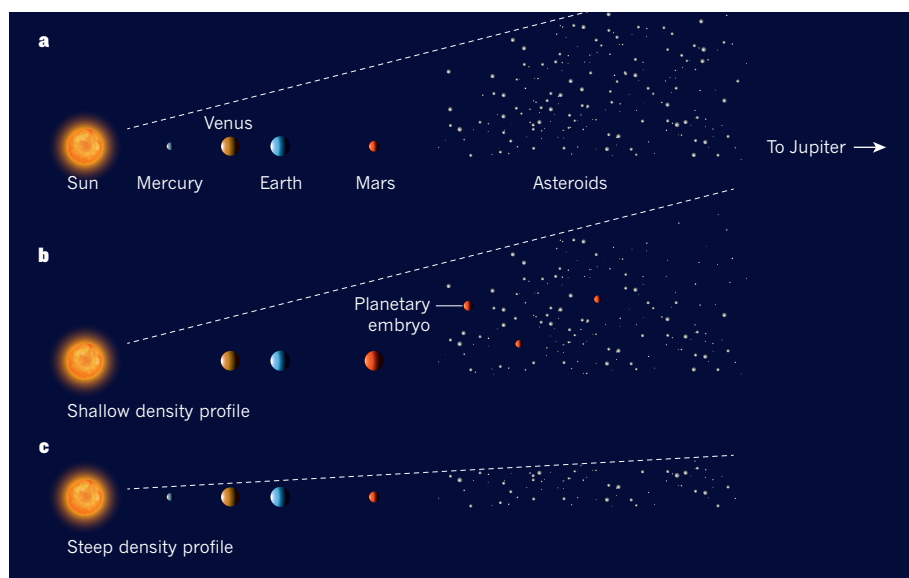
About 4.5 billion years ago, the Solar System formed in a disk of gas and dust particles that surrounded the newly born Sun<sup>1</sup>. The 'giant' planets (Jupiter, Saturn, Uranus and Neptune) formed first, within the few million years of the disk's lifetime. Closer to the Sun, the small, rocky 'terrestrial' planets (Mercury, Venus, Earth and Mars) took tens of million years to form, by collisions of numerous smaller objects generated in the disk. Myriad small bodies formed the asteroid belt between the orbits of Mars and Jupiter. Despite decades of attempts, no computational realization of standard formation theories has reproduced the mass and orbital distribution of both the terrestrial planets and the asteroids. Writing in the *Monthly Notices of the Royal Astronomical Society*, Izidoro *et al.*<sup>2</sup> show that this is not possible.

In the standard scenario of terrestrial-planet formation, planetary building blocks are segregated into two mass categories: a few tens of large planetary embryos, and several thousand small planetesimals from which embryos also grew while gas was still around. The embryos are roughly the size of Mars, whereas planetesimals are at most a few hundred kilometres across. All of these objects follow circular, co-planar orbits, and their number density decreases slowly with distance from the Sun. They are also gravitationally perturbed by the

giant planets, whose orbits have remained roughly unaltered since they began to form. As the system evolves, the strong gravitational pull that embryos receive from the giant planets and from each other deforms the embryos' orbits, which begin to cross. A cascade of collisions follows, forming planets as the embryos merge and collect planetesimals. Leftover planetesimals become asteroids.

Such simulations produce a small number of planets and a belt of leftovers that are broadly similar to the present Solar System, on reasonable timescales (see ref. 3, for example). Yet despite improvements, the detailed mass and orbital distribution of the inner Solar System are exceedingly difficult to reproduce. In the real Solar System, Venus and Earth are comparable in mass, and they orbit between the smaller Mercury and Mars (Fig. 1a). But standard models suffer from the 'Mars problem': in place of Mars, another planet forms that is comparable in size to Earth, and additional Mars-sized embryos can get stuck in the asteroid belt.

Although planets follow nearly circular, co-planar orbits, asteroid orbits are much more elliptical, and can be inclined to the ecliptic (Earth's mean orbital plane) by as much as 30°. Only part of this 'excitation' is explained by gravitational perturbations from the planets acting over the lifetime of the Solar System. If embryos had got stuck in the primordial belt, as suggested by standard simulations, this could have excited asteroid inclinations. But



**Figure 1 | The inner Solar System and outcomes of simulations of its formation.** In these cartoons, distances are roughly to scale, but the sizes of the planets relative to the Sun have been increased. **a**, The inner Solar System contains the terrestrial planets (Mercury, Venus, Earth and Mars) and the asteroid belt. Asteroid orbits can be inclined by up to  $30^\circ$  to the ecliptic (Earth's mean orbital plane). **b**, Simulations of the Solar System's formation in which material in the nascent system has a 'shallow' density profile (that is, planetary embryos can be found within 4 AU of the Sun; 1 AU is the distance from the Sun to Earth) typically result in an Earth-sized 'Mars' and asteroid orbits 'excited' at sizeable inclinations; Mercury is often absent. Embryos may get stuck in the asteroid belt for long times, eventually leading to a different belt from that now observed. **c**, Izidoro and colleagues<sup>2</sup> report simulations of steep initial density profiles (embryos within 2 AU of the Sun). The size and orbital distributions of the terrestrial planets are well reproduced in these simulations, but the asteroid belt remains nearly flat.

the embryos would have had to have decamped quickly, otherwise the asteroid belt would look very different today.

Building a huge 'Mars' and trapping massive embryos in the asteroid belt might not occur if most embryos are initially gathered within the orbit of Mars — that is, if a steep density profile develops in the early Solar System. This could occur if solid material in the gas disk accumulates near the Sun so that embryo formation is favoured in this region<sup>4</sup>. In their work, Izidoro *et al.* modelled the steep-profile scenario and compared it with other density profiles, but did not detail how this accumulation occurred.

Their main result is that, no matter what the density profile, it is impossible both to solve the Mars problem and to build a correctly structured asteroid belt (Fig. 1b, c). Steep density profiles reproduce the terrestrial planets fairly well, unlike shallow profiles. But asteroid excitation follows the opposite trend: steeper density profiles give much lower inclination excitations, because of the lack of embryos in the belt. Once the terrestrial planets have formed, leftover planetesimals cannot excite each other enough to produce the observed structure of the asteroid belt, because their gravity is weak. Perturbations from the giant planets cannot do the job either, even if they change their orbits abruptly later in the Solar System's evolution, as modern models predict (see ref. 5, for example). So the asteroid belt must have been both depleted of mass and excited before the terrestrial planets began to assemble.

Standard formation models don't consider the fact that giant planets can substantially change their orbits while forming in the disk. An intricate migration pattern of the giants has been reported<sup>6</sup> to produce a mass distribution that solves the Mars problem and generates an asteroid belt broadly similar to the observed one. As Izidoro *et al.* point out, this is the only known model that is compatible with their results, although it assumes a match between the growth and migration-time profiles of Jupiter and Saturn that has not been reproduced in simulations.

Planet-formation models have, in fact, been undergoing revision since it was realized that trillions of small pebbles in gas disks can spiral onto planetesimals as the pebbles drift towards the Sun, causing planetesimals to grow swiftly into embryos<sup>7</sup>. Pebble accretion, combined with gravitational self-stirring, has been shown to produce the giant planets within the short lifetime of the disk<sup>8</sup> — the first time that this has been achieved in a model. More-recent work<sup>9</sup> suggests that the same process might explain the structure of the inner Solar System, without the need for giant-planet migration. However, this accretion model also depends sensitively on largely unknown physical properties of both the gas disk and the growing bodies.

Izidoro *et al.* do not offer a final model of terrestrial-planet formation. But their work convincingly demonstrates that standard models cannot satisfy major constraints on the process, the toughest of which is set by asteroids.



## 50 Years Ago

Fortunately, our distant ancestors appear to have had a mania for making lists. Some of these lists are in a sense the beginnings of history, just as others, which perform a preliminary work of classification, are in a sense the beginning of the natural sciences. It was when they put together their lists of successive kings or priests that these ancient peoples acquired their first impressions of the tremendous stretch of time behind them. The ancient Greeks had very defective lists and thought that only a comparatively short period separated them from the age when the gods had walked and sported about on the Earth. But some of them learned about the huge lists of Egyptian priests and came to realize that there had been thousands of years of human history before their day. If the whole of this is taken together, it involves a change in what it means to have an existence in time—a time which stretches behind and before.

From *Nature* 11 December 1965

## 100 Years Ago

*First Aid in the Laboratory and Workshop* by A. A. Eldridge and Dr. H. V. A. Briscoe

The authors of this little book, who have been in charge of first aid organisation in chemical and physical laboratories, have found that the ordinary text-books devote too much space to serious fractures and other injuries, but give little information regarding ordinary accidents, such as are apt to occur in laboratories and workshops, for instance, burns produced by chemicals, eye injuries, shocks produced by electric currents, and poisoning. They have therefore written this pamphlet to meet this need.

From *Nature* 9 December 1915



Even if their simulations were refined, it is unlikely that this general result would change. Planetary scientists should now focus on whether the intricate structure of the inner Solar System can be adequately explained by non-standard accretion models, or whether it simply represents the heritage of a preceding phase of extensive giant-planet migration<sup>6</sup>. ■

**Kleomenis Tsiganis** is in the Department of Physics, Aristotle University of Thessaloniki,

GR 54124 Thessaloniki, Greece.  
e-mail: [tsiganis@auth.gr](mailto:tsiganis@auth.gr)

1. Armitage, P. J. *Astrophysics of Planet Formation* (Cambridge Univ. Press, 2010).
2. Izidoro, A., Raymond, S. N., Morbidelli, A. & Winter, O. C. *Mon. Not. R. Astron. Soc.* **453**, 3619–3634 (2015).
3. Raymond, S. N., Quinn, T. & Lunine, J. I. *Icarus* **183**, 265–282 (2006).
4. Ogihara, M., Kobayashi, H., Inutsuka, S. & Suzuki, T. K. *Astron. Astrophys.* **579**, A65 (2015).
5. Morbidelli, A., Levison, H. F., Tsiganis, K. & Gomes, R. *Nature* **435**, 462–465 (2005).

6. Walsh, K. J., Morbidelli, A., Raymond, S. N., O'Brien, D. P. & Mandell, A. M. *Nature* **475**, 206–209 (2011).
7. Lambrechts, M. & Johansen, A. *Astron. Astrophys.* **544**, A32 (2012).
8. Levison, H. F., Kretke, K. A. & Duncan, M. *Nature* **524**, 322–324 (2015).
9. Levison, H. F., Kretke, K. A., Walsh, K. J. & Bottke, W. F. *Proc. Natl Acad. Sci. USA* **112**, 14180–14185 (2015).

**The author declares competing interests:** See online article for details.

**This article was published online on 2 December 2015.**

## STRUCTURAL BIOLOGY

# A transcriptional specialist resolved

**Three structures of the enzyme RNA polymerase III, which is responsible for the synthesis of abundant short RNAs, reveal the specializations that make it an adept terminator and reinitiator of transcription. [SEE ARTICLE P.231](#)**

RICHARD J. MARAIA & KESHAB RIJAL

**R**NA polymerase enzymes have no small job. Present in every cell, they are essential for transcribing DNA into RNA. Bacteria and archaea each use a single RNA polymerase (Pol), whereas all other organisms (called eukaryotes) have three specialized Pools<sup>1</sup> — Pol I and Pol II, which synthesize different types of long RNA, and Pol III, which makes short RNAs. In this issue, Hoffmann *et al.*<sup>2</sup> (page 231) report three structures of Pol III, all at near-atomic resolution. These structures allow the authors to make comparisons with existing Pol I and Pol II structures, and to suggest how Pol III terminates and reinitiates transcription. Moreover, their work completes the set of five Pol structures.

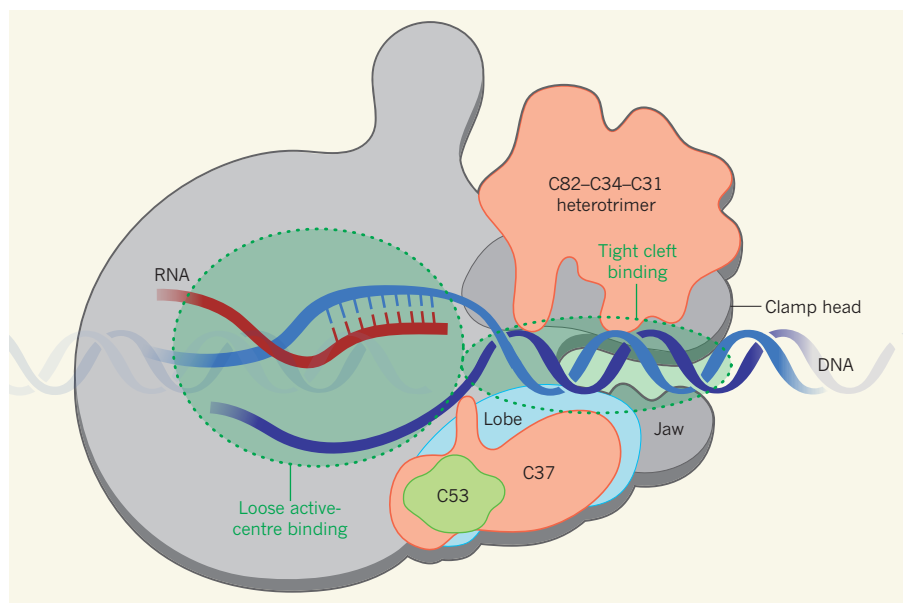
Pol III produces a huge supply of short structural RNAs that collectively outnumber all other RNAs in the cell, and the enzyme must therefore initiate and terminate transcription more frequently than Pools I or II (ref. 3). Pol III is adapted for this role — it is the largest of the three eukaryotic Pol enzymes, with 17 subunits, some of which are Pol III-specific relatives of transcription factors that transiently associate with Pol II during initiation. The stable association of these subunits with Pol III enables efficient initiation of transcription and enzyme recycling<sup>4</sup>. The Pol III-specific subunits are organized into two subcomplexes, a C82–C34–C31 heterotrimer and a C53–C37 heterodimer, the latter of which is also involved in transcription termination<sup>5</sup>. But precisely how these subunits are positioned so as to contribute to the specialized functions of Pol III has remained unclear.

During transcriptional elongation of the

nascent RNA strand, Pol III acts in a closed-clamp conformation, in which a cleft in the enzyme exerts a tight grip on the DNA awaiting transcription. Hoffmann *et al.* used cryo-electron microscopy to determine the structures of apoenzyme Pol III (the enzyme minus the DNA and RNA) from the brewer's

yeast *Saccharomyces cerevisiae* in both closed- and open-clamp conformations, achieving resolutions of 4.6 and 4.7 ångströms, respectively. They also resolved the structure of the enzyme complexed with RNA and DNA during active elongation, to 3.9 Å.

Hoffmann and colleagues' structures reveal that the cleft of Pol III grips DNA more tightly than that of other Pools, in part because the clamp-head domain in subunit C160 is larger than in the other enzymes. The C82–C34–C31 heterotrimer packs onto this clamp head through C82–C160 interactions, and extensions of C82 wedge the DNA in the cleft at around 15 and 7 base pairs away from the active centre, where transcription occurs. Several sections on the other side of the heterotrimer face and are close to a stalk structure, which is common in Pools and is involved in initiation in Pol III. The stalk adopts different orientations in the two apoenzymes, possibly



**Figure 1 | Function encoded in structure.** Hoffmann *et al.*<sup>2</sup> resolved three structures of the enzyme RNA polymerase (Pol) III, which transcribes short RNAs. Pol III contains specialized subunits — the C82–C34–C31 heterotrimer and the C53–C37 heterodimer — that enable it to terminate and reinitiate transcription. The authors find that the heterotrimer packs onto the enzyme's clamp-head domain, and extends helices that help to hold the DNA awaiting transcription in a tight grip in the structure's cleft. The heterodimer connects to the lobe and jaw domains on the lower side of the cleft. By contrast, DNA in the active centre (where transcription occurs) is held loosely. During termination, the untranscribed DNA strand, which contains terminator signal sequences, makes contact with C37. C37 might transmit termination signals to the lower side of the cleft, releasing the downstream DNA. The authors propose that the loose grip in the active centre then enables the DNA to easily separate from the enzyme. For simplicity, not all Pol III subunits are shown.

reflecting different stages of transcription.

The structures provide several clues to how Pol III is adapted for termination and initiation. The genes transcribed by Pol III end with a short string of thymine (T) bases called the terminator sequence<sup>6</sup>; its transcription produces an unstable RNA–DNA hybrid that causes Pol III to dissociate from DNA<sup>6</sup>. This process is aided by terminator T sequences in the complementary, untranscribed DNA strand. Termination and subsequent reinitiation are known<sup>5</sup> to involve the collective activities of the C53–C37 heterodimer and another subunit required for efficient termination and reinitiation cycles, C11.

The authors find that C53–C37 is connected to C11 and attached to lobe and jaw domains (named after their shape), which lie across the cleft from the clamp head. C37 connects to DNA in the cleft through a helical extension, then continues to the active centre and makes contact with untranscribed sequences from the complementary DNA. This structure suggests a model in which C37 transmits the terminator signal to the cleft through a mechanism involving C11 (refs 7, 8), to trigger opening of the clamp. Such a set-up might ensure that the downstream DNA is released in sync with termination (Fig. 1).

The C11 subunit contains two terminal domains, which are related to proteins called Rpb9 and TFIIS (ref. 9) that interact with Pol II. The TFIIS-related carboxy-terminal domain of C11 mediates cleavage of the end of the RNA in the active centre when the enzyme pauses (a prerequisite for termination). Hoffmann *et al.* report that, in the apoenzyme, this region is stored close to an entry gate to the active centre called the funnel pore. It is tethered by a linker to C11's amino-terminal domain, which is anchored by multiple contacts to C37, the lobe and the jaw, in agreement with an earlier study<sup>10</sup>. This set-up provides a firm anchor from which the C-terminal domain can swing on the linker, accessing the active centre during pausing, termination and, perhaps, reinitiation. Indeed, the authors could not map the C-terminal domain in the complexed structure, presumably because of its dynamic interactions with the active centre.

Of particular interest is the architecture of the active centre. Hoffmann and colleagues found that this centre grips the RNA–DNA hybrid only loosely, and they suggest that the unusually tight downstream cleft compensates for this during elongation. Because 'melting' of the intrinsically weak hybrid into single strands is a key determinant of termination<sup>11</sup>, this loose grip probably allows Pol III to readily release the RNA from the active centre on cue.

These structures complete the triad of Pols that evolved from the bacterial enzyme, and provide explanations for the specialized properties that enable division of labour among the eukaryotic Pols. The three structures of Pol III give us much more information than one can.

However, each represents only one state of a dynamic process. There is more to be learnt — for instance, large parts of the heterotrimer subunits C34 and C31 are yet to be resolved. Much of the heterodimer subunit C53 also remains unresolved, as do the interactions between the complementary DNA strand and C37, the lobe and the jaw. A better understanding of these subunits should provide further insights into the mechanism by which termination occurs. Moreover, they might provide clues to how Pol III is reset for initiation.

Hoffmann and colleagues' structures are in agreement with much existing physical, biochemical and genetic data<sup>5–9,12,13</sup>. They are also excellent substrates for mutational and mechanistic studies that could cast light on how cancers that depend on Pol III for growth might be attacked<sup>14</sup>. ■

**Richard J. Maraia and Keshab Rijal** are in the Intramural Research Program, Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda,

Maryland 20892–2425, USA.  
e-mails: maraiar@mail.nih.gov;  
keshab.rijal@nih.gov

1. Werner, F. & Grohmann, D. *Nature Rev. Microbiol.* **9**, 85–98 (2011).
2. Hoffmann, N. A. *et al. Nature* **528**, 231–236 (2015).
3. Moir, R. D. & Willis, I. M. *Biochim. Biophys. Acta* **1829**, 361–375 (2013).
4. Arimbasseri, A. G., Rijal, K. & Maraia, R. J. *Transcription* **5**, e27639 (2014).
5. Landrieux, E. *et al. EMBO J.* **25**, 118–128 (2006).
6. Arimbasseri, A. G. & Maraia, R. J. *Mol. Cell* **58**, 1124–1132 (2015).
7. Iben, J. R. *et al. Nucleic Acids Res.* **39**, 6100–6113 (2011).
8. Rijal, K. & Maraia, R. J. *Nucleic Acids Res.* **41**, 139–155 (2013).
9. Chédin, S., Riva, M., Schultz, P., Sentenac, A. & Carles, C. *Genes Dev.* **12**, 3857–3871 (1998).
10. Fernández-Tornero, C. *et al. EMBO J.* **29**, 3762–3772 (2010).
11. Martin, F. H. & Tinoco, I. Jr *Nucleic Acids Res.* **8**, 2295–2300 (1980).
12. Wu, C.-C., Lin, Y.-C. & Chen, H.-T. *Mol. Cell. Biol.* **31**, 2715–2728 (2011).
13. Khoo, S.-K., Wu, C.-C., Lin, Y.-C., Lee, J.-C. & Chen, H.-T. *Mol. Cell. Biol.* **34**, 551–559 (2014).
14. Johnson, S. A. S., Dubeau, L. & Johnson, D. L. *J. Biol. Chem.* **283**, 19184–19191 (2008).

This article was published online on 25 November 2015.

## IMMUNOLOGY

# In the right place at the right time

**Regulatory T cells help to prevent autoimmune responses. A new imaging technique reveals that activation of these cells requires clustering with self-reactive effector T cells and sensing of the signalling protein interleukin-2. [SEE ARTICLE P.225](#)**

**ESTEBAN CARRIZOSA  
& THORSTEN R. MEMPEL**

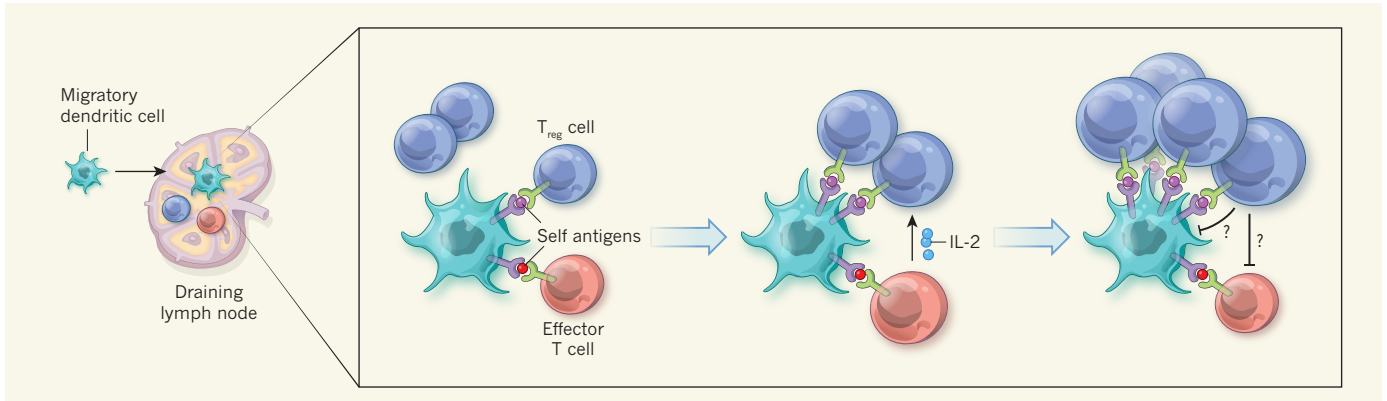
**T**he T cells of our immune system produce receptor proteins that specifically bind to molecular structures called antigens. Each T cell expresses one type of receptor, and the immense diversity of the T-cell population ensures that our immune system can detect and respond in a highly specific manner to almost any pathogen we encounter. But a fundamental problem of this sensing strategy is that it also generates T cells that bear self-reactive receptors, which can trigger dangerous autoimmune attacks. Specialized regulatory T cells are needed to keep such autoreactive effector T cells in check, but exactly how they achieve this task is incompletely understood. On page 225 of this issue, Liu *et al.*<sup>1</sup> provide information on how the spatial organization of regulatory and effector T cells helps to maintain immune homeostasis.

For many years, conventional wisdom held that, as T cells develop in the thymus and become equipped with T-cell receptors (TCRs) of unpredictable specificity, they are subjected

to a negative-selection process in which cells bearing receptors that react strongly with self antigens are removed. Recently, however, it has become evident<sup>2</sup> that numerous autoreactive T cells continually leave the thymus and enter secondary lymphoid organs, such as the spleen and lymph nodes, where immune responses are initiated. Keeping these autoreactive T cells in check is achieved in part by regulatory T (T<sub>reg</sub>) cells, which mostly bear TCRs that also bind self antigens. When T<sub>reg</sub> cells lose their function, potentially lethal autoimmune diseases develop in both mice and humans<sup>3</sup>.

What is it that guides T<sub>reg</sub> cells to focus their immune-regulatory activities on autoreactive effector T cells in the vast expanse of the secondary lymphoid organs, where they are outnumbered 10 to 1 by their non-regulatory counterparts? Liu *et al.* observed in healthy mice that most T<sub>reg</sub> cells are evenly dispersed in the areas of lymphoid tissues that are densely packed with T cells, but some small clusters of T<sub>reg</sub> cells form in the lymph nodes' superficial paracortex. This region has the highest density of migratory dendritic cells (DCs) — immune cells that survey tissues and take up both self





**Figure 1 | Regulatory T-cell control of autoreactive T cells.** Some types of the immune system's dendritic cells (DCs) collect molecular samples of our tissue environments and migrate to draining lymph nodes. Here, they present their samples to T cells, which express unique surface receptors that recognize specific molecular structures (antigens). Some of these T-cell receptors will bind self antigens; if such binding occurs and the T cell is activated, an autoimmune response could ensue. However, a specialized class of T cell — regulatory T ( $T_{reg}$ ) cells, which also recognize self antigens — can be activated

to suppress these effector T-cell responses. Liu *et al.*<sup>1</sup> suggest that  $T_{reg}$  and autoreactive effector T cells encounter their specific antigens on migratory DCs, leading to initial activation and clustering of  $T_{reg}$  cells close to effector T cells that produce the cell-signalling molecule IL-2. This cytokine then fuels further  $T_{reg}$ -cell proliferation and/or function, and the fully activated  $T_{reg}$  cells eventually suppress effector T-cell activation, possibly by acting directly on the effector T cells or indirectly by suppressing the stimulatory capacity of the DCs.

and foreign antigens before migrating through the lymphatic system to the tissue-draining lymph nodes to present these antigens to T cells.

The observation that clustered  $T_{reg}$  cells are in preferential contact with migratory DCs even in the absence of foreign antigens suggests that the clustering is driven by the recognition of self antigens (Fig. 1). Indeed, the authors observed that  $T_{reg}$  cells lost the ability to form clusters when their TCRs were deleted. Two previous studies<sup>4,5</sup> showed that TCR signalling is required for mature  $T_{reg}$  cells to acquire and sustain their full suppressive activity, which includes elevated expression of regulatory molecules such as CD73 and CTLA4. Intriguingly, clustered  $T_{reg}$  cells also expressed the largest quantities of these molecules, and this property was lost after TCR deletion, indicating that TCR triggering not only positions  $T_{reg}$  cells in clusters, but also increases their suppressive activity.

In addition to these TCR-dependent events, Liu *et al.* noted that the transcription factor STAT5 was phosphorylated (indicating that it was activated) in most clustering  $T_{reg}$  cells. This phosphorylation was driven by a cytokine (a secreted cell-signalling protein) called interleukin (IL)-2, which was produced by effector T cells at the centre of  $T_{reg}$ -cell clusters. When the researchers treated their mice with an antibody that inhibits IL-2 function, they observed increased activation of autoreactive effector T cells, indicating a loss of suppression by  $T_{reg}$  cells. This finding also suggests that the initial activation-driven proliferation of effector T cells does not require the IL-2 that they themselves produce, as previously reported<sup>6</sup>. Instead, the cytokine primarily activates a negative feedback loop involving IL-2-responsive  $T_{reg}$  cells that are recruited to sites of autoreactive effector T-cell activation, to stop these dangerous responses in their tracks (Fig. 1).

Liu and colleagues' study suggests that the activation of autoreactive effector T cells is a surprisingly frequent event that seemingly brings us to the verge of autoimmune disease, but that rapid recruitment of  $T_{reg}$  cells into functional niches containing self-antigen-presenting DCs ensures that these self antigens can be tolerated. Further research is needed to determine whether clustering is essential for suppression or merely a side effect of the TCR-driven  $T_{reg}$ -cell activation process.

The precise mechanism by which  $T_{reg}$ -cell clusters are formed is also still unclear. The most straightforward explanation is that they arise when  $T_{reg}$  cells stabilize their interactions with those DCs that were activated either through previous interactions with autoreactive T effector cells or by other means (which could have triggered the autoreactive effector T-cell response in the first place). Another, but not mutually exclusive, possibility is that TCR signalling in  $T_{reg}$  cells renders them responsive to chemoattractant molecules produced by activated DCs, rather like the spatial organization that allows one class of effector T cell ( $CD4^+$  T cells) to provide 'help' to another ( $CD8^+$  T cells)<sup>7</sup>. These considerations also raise the question of whether clustering  $T_{reg}$  and autoreactive effector T cells respond to related self antigens, and whether they may encounter them on different DCs or must interact with the same DC.

It will also be interesting to explore the fates of autoreactive effector T cells once they are brought under control, and to investigate whether this interplay between  $T_{reg}$  and effector T cells applies only in lymphoid tissues, or also at effector sites of immune responses, where regulatory function requires that both  $T_{reg}$  cells and effector T cells re-encounter their specific antigens<sup>8</sup>. Furthermore, how can the barrier to complete T-cell activation described in this study be overcome during appropriate

responses against pathogens, or break down where autoimmunity occurs?

Finally, it is appropriate to point out the tremendous potential of the integrative tissue-imaging technique used in this study, which the authors have named histocytometry<sup>9</sup>. In contrast to other multiplexed single-cell-analysis techniques, this method records the position of multiple cell types in intact tissues, along with quantitative information on their activation states and gene-expression patterns. As illustrated here, this can reveal minor functional populations that participate in local cellular networks but would probably be overlooked in any type of global analysis. In times of increasingly fine-grained molecular analyses and definitions of subsets of immune cells, detailed information on their micro-environmental contexts promises to eventually produce a true systems-level understanding of the immune system. ■

**Esteban Carrizosa and Thorsten R. Mempel** are at Harvard Medical School and the Center for Immunology and Inflammatory Diseases, Massachusetts General Hospital, Boston, Massachusetts 02114, USA.

e-mail: tmempel@mgh.harvard.edu

1. Liu, Z. *et al.* *Nature* **528**, 225–230 (2015).
2. Hogquist, K. A. & Jameson, S. C. *Nature Immunol.* **15**, 815–823 (2014).
3. Benoist, C. & Mathis, D. *Cold Spring Harb. Perspect. Biol.* **4**, a007021 (2012).
4. Vahl, J. C. *et al.* *Immunity* **41**, 722–736 (2014).
5. Levine, A. G., Arvey, A., Jin, W. & Rudensky, A. Y. *Nature Immunol.* **15**, 1070–1078 (2014).
6. O'Gorman, W. E. *et al.* *J. Immunol.* **183**, 332–339 (2009).
7. Castellino, F. *et al.* *Nature* **440**, 890–895 (2006).
8. Bauer, C. A. *et al.* *J. Clin. Invest.* **124**, 2425–2440 (2014).
9. Gerner, M. Y., Kastenmüller, W., Ifrim, I., Kabat, J. & Germain, R. N. *Immunity* **37**, 364–376 (2012).

This article was published online on 25 November 2015.

# Undecidability of the spectral gap

Toby S. Cubitt<sup>1,2</sup>, David Perez-Garcia<sup>3,4</sup> & Michael M. Wolf<sup>5</sup>

**The spectral gap—the energy difference between the ground state and first excited state of a system—is central to quantum many-body physics. Many challenging open problems, such as the Haldane conjecture, the question of the existence of gapped topological spin liquid phases, and the Yang–Mills gap conjecture, concern spectral gaps. These and other problems are particular cases of the general spectral gap problem: given the Hamiltonian of a quantum many-body system, is it gapped or gapless? Here we prove that this is an undecidable problem. Specifically, we construct families of quantum spin systems on a two-dimensional lattice with translationally invariant, nearest-neighbour interactions, for which the spectral gap problem is undecidable. This result extends to undecidability of other low-energy properties, such as the existence of algebraically decaying ground-state correlations. The proof combines Hamiltonian complexity techniques with aperiodic tilings, to construct a Hamiltonian whose ground state encodes the evolution of a quantum phase-estimation algorithm followed by a universal Turing machine. The spectral gap depends on the outcome of the corresponding ‘halting problem’. Our result implies that there exists no algorithm to determine whether an arbitrary model is gapped or gapless, and that there exist models for which the presence or absence of a spectral gap is independent of the axioms of mathematics.**

The spectral gap is one of the most important physical properties of a quantum many-body system, determining much of its low-energy physics. Gapped systems exhibit non-critical behaviour (for example, massive excitations and short-range correlations), whereas phase transitions occur when the spectral gap vanishes and the system exhibits critical behaviour (for example, massless excitations and long-range correlations). Many seminal results in condensed matter theory prove that specific systems are gapped or gapless, for example, that the Heisenberg chain is gapless for half-integer spin<sup>1</sup> (later extended to higher dimensions<sup>2</sup>), or that the 1D AKLT (Affleck–Kennedy–Lieb–Tasaki) model is gapped<sup>3</sup>. Similarly, many famous and long-standing open problems in theoretical physics concern the presence or absence of a spectral gap. A paradigmatic example is the antiferromagnetic Heisenberg model in 1D with integer spins. The ‘Haldane conjecture’ that this model is gapped, first formulated in 1983<sup>4</sup>, has yet to be rigorously proven despite strong supporting numerical evidence<sup>5</sup>. The same question in the case of 2D non-bipartite lattices such as the kagome lattice was posed in 1973<sup>6</sup>. Numerical evidence<sup>7</sup> strongly indicates that these systems may be topological spin liquids. This problem has attracted substantial attention<sup>8</sup> because materials such as herbertsmithite<sup>9</sup> have emerged whose interactions are well-approximated by the Heisenberg coupling. The presence of a spectral gap in these models remains one of the main unsolved questions concerning the long-sought topological spin liquid phase. In the related setting of quantum field theory, one of the most notorious open problems again concerns a spectral gap—the Yang–Mills mass gap problem<sup>10</sup>. Proving the existence of a gap in Yang–Mills theory could provide a full explanation of the phenomenon of quark confinement. Although there is strong supporting evidence of such a gap from numerical lattice quantum chromodynamics computations<sup>11</sup>, the problem remains open.

All of these problems are specific instances of the general spectral gap problem: given a quantum many-body Hamiltonian, is the system it describes gapped or gapless? Our main result is to prove that the spectral gap problem is undecidable in general. This involves more than merely showing that the problem is computationally or mathematically

hard. Although one may be able to solve the spectral gap problem in specific cases, our result implies that it is, in general, logically impossible to determine whether a system is gapped or gapless. This statement has two meanings, and we prove both.

(1) The spectral gap problem is algorithmically undecidable: there cannot exist any algorithm that, given a description of the local interactions, determines whether the resultant model is gapped or gapless. This is the same sense in which the halting problem is undecidable<sup>12</sup>.

(2) The spectral gap problem is axiomatically independent: given any consistent recursive axiomatization of mathematics, there exist particular quantum many-body Hamiltonians for which the presence or absence of the spectral gap is not determined by these axioms. This is the form of undecidability encountered in Gödel’s incompleteness theorem<sup>13</sup>.

## Precise statement of results

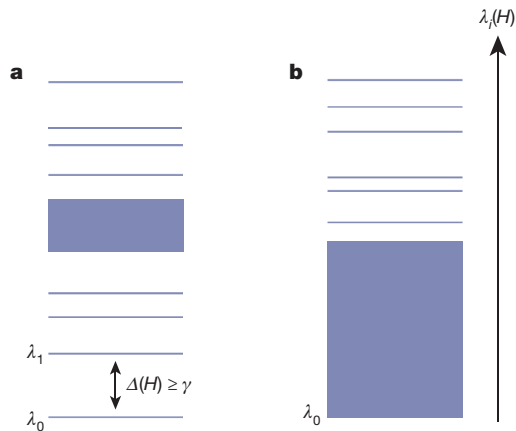
It is important to be precise in what we mean by the spectral gap problem. To this end, we must first specify the systems we are considering. Because we are proving undecidability, the simpler the system, the stronger the result. We restrict ourselves to nearest-neighbour, translationally invariant spin lattice models on a 2D square lattice of size  $L \times L$  (which we later take to  $\infty$ ), with local Hilbert space dimension  $d$ . Any such Hamiltonian  $H_L$  is completely specified by at most three finite-dimensional Hermitian matrices describing the local interactions of the system: two  $d^2 \times d^2$  matrices  $h_{\text{row}}$  and  $h_{\text{col}}$  that specify the interactions along the rows and columns of the lattice, and a  $d \times d$  matrix  $h_1$  that specifies any on-site interaction. All matrix elements will be algebraic numbers, and we normalize the interaction strength such that  $\max\{\|h_{\text{row}}\|, \|h_{\text{col}}\|, \|h_1\|\} = 1$ .

We must also be precise in what we mean by ‘gapped’ and ‘gapless’ (see Fig. 1). Because quantum phase transitions occur in the thermodynamic limit of arbitrarily large system size, we are interested in the spectral gap  $\Delta(H_L) = \lambda_1(H_L) - \lambda_0(H_L)$  as the system size  $L \rightarrow \infty$  (where  $\lambda_0$  and  $\lambda_1$  are the eigenvalues of  $H_L$  with the smallest and second-smallest magnitude). We take ‘gapped’ to mean that the system has a unique ground state and a constant lower bound on the spectral gap:

<sup>1</sup>Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK. <sup>2</sup>DAMTP, University of Cambridge, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK. <sup>3</sup>Departamento de Análisis Matemático and IMI, Facultad de CC Matemáticas, Universidad Complutense de Madrid, Plaza de Ciencias 3, 28040 Madrid, Spain.

<sup>4</sup>ICMAT, C/Nicolás Cabrera, Campus de Cantoblanco, 28049 Madrid, Spain. <sup>5</sup>Department of Mathematics, Technische Universität München, 85748 Garching, Germany.





**Figure 1 | Gapped and gapless systems.** **a**, A gapped system has a unique ground state  $\lambda_0(H)$  and a constant lower-bound  $\gamma$  on the spectral gap  $\Delta(H) = \lambda_1 - \lambda_0$  in the thermodynamic limit. **b**, A gapless system has continuous spectrum  $\lambda_i(H)$  above the ground state in the thermodynamic limit.

$\Delta(H_L) \geq \gamma > 0$  for all sufficiently large  $L$ . We take ‘gapless’ to mean the system has continuous spectrum above the ground state in the thermodynamic limit.

Here gapped is not the negation of gapless; there are systems that fall into neither category. We adopt such strong definitions to deliberately exclude ambiguous cases, such as systems with degenerate ground states. A Hamiltonian that is gapped or gapless according to the above definitions is recognized as such throughout the literature. We show that the spectral gap problem is undecidable even given that the Hamiltonian either has a unique ground state and a spectral gap of magnitude one, or has continuous spectrum above the ground state.

We prove this by showing that the halting problem for Turing machines can be encoded in the spectral gap problem, implying that the latter is at least as hard as the former. A Turing machine is a simple, abstract model of computation in which a head reads and writes symbols from some finite alphabet on an infinite tape and moves left or right, following a finite set of rules. The halting problem asks: given an initial input written on the tape, does the Turing machine halt? Turing proved that this problem is undecidable<sup>12</sup>; we relate it to the spectral gap problem in the following way.

### Theorem 1

We can explicitly construct a dimension  $d$ ,  $d^2 \times d^2$  matrices  $A$ ,  $B$ ,  $C$  and  $D$ , and a rational number  $\beta > 0$ , which can be chosen to be as small as desired, such that

- (i)  $A$  is Hermitian, with matrix elements in  $\mathbb{Z} + \beta\mathbb{Z} + \frac{\beta}{\sqrt{2}}\mathbb{Z}$ ;
- (ii)  $B$  and  $C$  have integer matrix elements; and
- (iii)  $D$  is Hermitian, with matrix elements in  $\{0, 1, \beta\}$ .

For each positive integer  $n$ , define the local interactions of a translationally invariant, nearest-neighbour Hamiltonian  $H(n)$  on a 2D square lattice as

$$\begin{aligned} h_1(n) &= \alpha(n)II \\ h_{\text{row}} &= D \\ h_{\text{col}} &= A + \beta(e^{i\pi\varphi(n)}B + e^{-i\pi\varphi(n)}B^\dagger + e^{i\pi 2^{-|\varphi(n)|}}C + e^{-i\pi 2^{-|\varphi(n)|}}C^\dagger) \end{aligned}$$

where  $\varphi(n) = n/2^{|\varphi(n)|-1}$  is the rational number whose binary fraction expansion contains the binary digits of  $n$  after the decimal point,  $|\varphi(n)|$  denotes the number of digits in this expansion,  $\alpha(n) \leq \beta$  is an algebraic number that is computable from  $n$ ,  $II$  is a projector and the daggers denote Hermitian conjugation. Then

- (i) the local interaction strength is  $\leq 1$  (that is,  $\|h_1(n)\|, \|h_{\text{row}}\|, \|h_{\text{col}}(n)\| \leq 1$ );
- (ii) if the universal Turing machine halts on input  $n$ , the Hamiltonian  $H(n)$  is gapped with  $\gamma \geq 1$ ; and

- (iii) if the universal Turing machine does not halt on input  $n$ , the Hamiltonian  $H(n)$  is gapless (that is, has continuous spectrum).

Theorem 1 implies that the spectral gap problem is algorithmically undecidable because the halting problem is. By a standard argument<sup>14</sup> algorithmic undecidability also implies axiomatic independence. Both forms of undecidability extend to other low-temperature properties of quantum systems, such as critical correlations in the ground state. In fact, our method allows us to prove undecidability of any physical property that distinguishes a Hamiltonian from a gapped system with unique, product ground state.

### Hamiltonian construction

We first relate undecidability of the spectral gap to undecidability of another important physical quantity, the ground state energy density, which, for a 2D lattice, is given by  $E_g = \lim_{L \rightarrow \infty} [\lambda_0(H_L)/L^2]$ . We then transform the halting problem into a question about ground state energy densities.

Reducing the ground state energy density problem to the spectral gap problem requires two ingredients.

(1) It requires a translationally invariant Hamiltonian  $H_u(\varphi)$  on a 2D square lattice with local interactions  $h_u(\varphi)$ , whose ground state energy density is either strictly positive or tends to zero from below in the thermodynamic limit, depending on the value of an external parameter  $\varphi$ ; however, determining which case holds should be undecidable. Constructing such a Hamiltonian constitutes the main technical work of our result. (These properties of  $H_u(\varphi)$  are unaffected if we multiply  $h_u(\varphi)$  by an arbitrary fixed rational number  $\beta$ , no matter how small.)

(2) It requires a gapless Hamiltonian  $H_d$  with translationally invariant local interactions  $h_d$  and a ground state energy of zero. (Recall that by ‘gapless’ we mean continuous spectrum above the ground state, not merely a vanishing spectral gap.) There are many well-known examples of such Hamiltonians, for example, that associated with the critical XY model<sup>1</sup>.

Given Hamiltonians with these properties, we construct a new translationally invariant Hamiltonian, with local interactions  $h(\varphi)$ , that is gapped or gapless depending on the value of  $\varphi$ . The local Hilbert space of  $h(\varphi)$  is the tensor product of those of  $h_u$  and  $h_d$  together with one additional energy level:  $\mathcal{H} = |0\rangle \oplus \mathcal{H}_u \otimes \mathcal{H}_d$ . We take the interaction  $h^{(i,j)}$  between nearest-neighbour sites  $i$  and  $j$  to be

$$\begin{aligned} h(\varphi)^{(i,j)} &= |0\rangle\langle 0|^{(i)} \otimes (\mathbb{1} - |0\rangle\langle 0|)^{(j)} + h_u^{(i,j)}(\varphi) \otimes \mathbb{1}_d^{(i,j)} \\ &\quad + \mathbb{1}_u^{(i,j)} \otimes h_d^{(i,j)} \end{aligned} \quad (1)$$

The spectrum of the new Hamiltonian  $H$  is

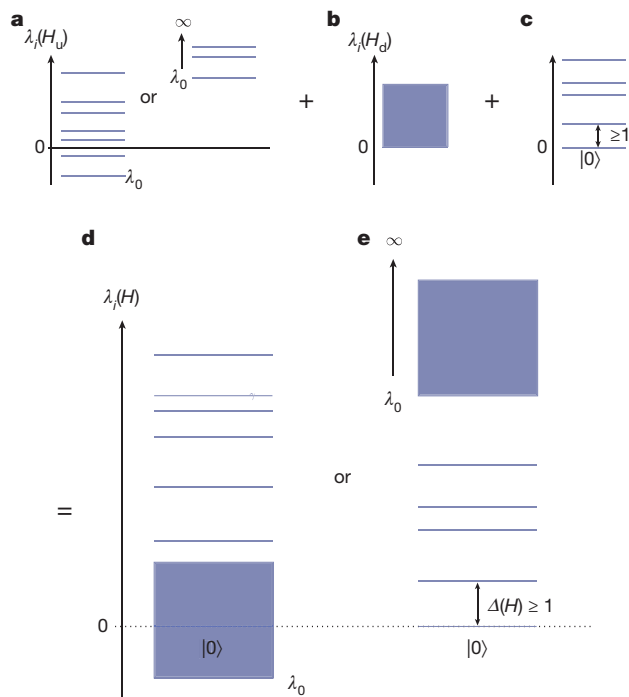
$$\text{spec}H = \{0\} \cup \{\text{spec}H_u(\varphi) + \text{spec}H_d\} \cup S \quad (2)$$

with  $S \geq 1$  (see Supplementary Information for details). Recalling that we chose  $H_d$  to be gapless, we see immediately from equation (2) that if the ground state energy density of  $H_u$  tends to zero from below (so that  $\lambda_0(H_u) < 0$ ), then  $H(\varphi)$  is gapless; if  $H_u$  has a strictly positive ground state energy density (so that  $\lambda_0(H_u)$  diverges to  $+\infty$ ), then it has a spectral gap  $\geq 1$ , as required (see Fig. 2).

This construction is rather general: by choosing different  $h_d$ , we obtain undecidability of any physical property that distinguishes a Hamiltonian from a gapped system with a unique product ground state.

### Encoding computation in ground states

To construct the Hamiltonian  $H_u(\varphi)$ , we encode the halting problem into the local interactions  $h_u(\varphi)$  of the Hamiltonian. The halting problem concerns the dynamics of a classical system—a Turing machine. To relate it to the ground state energy density—a static property of a quantum system—we construct a Hamiltonian whose ground state encodes the entire history of the computation carried out by the Turing



**Figure 2 | Relating ground state energy density to spectral gap.** **a–c,** To relate ground state energy density and spectral gap, we need a Hamiltonian  $H_u(\varphi)$  whose ground state energy density is either strictly positive or tends to zero from below in the thermodynamic limit, but determining which is undecidable (**a**), and a gapless Hamiltonian  $H_d$  with a ground state energy of zero (**b**). We combine  $H_u(\varphi)$  and  $H_d$  to form a new local interaction,  $h(\varphi)$ , in such a way that  $H(\varphi)$  has an additional non-degenerate zero-energy eigenstate  $|0\rangle$  (**c**), and that the continuous spectrum of  $H_d$  is shifted immediately above the ground state energy of  $H_u$ . **d,** If the ground state energy density of  $H_u(\varphi)$  tends to zero from below, then its ground state energy in the thermodynamic limit must be  $\leq 0$ , and  $H(\varphi)$  is gapless. **e,** Alternatively, if the ground state energy density of  $H_u(\varphi)$  is strictly positive, then its ground state energy in the thermodynamic limit must diverge to  $+\infty$ , and  $H(\varphi)$  is gapped with gap  $\Delta(H) \geq 1$ .

machine in superposition<sup>15</sup>: if the state of the computation at time  $t$  is represented by the state vector  $|\psi_t\rangle$ , and the computation runs until time  $T$ , then the ground state is the so-called ‘computational history state’  $\frac{1}{\sqrt{T}} \sum_{t=0}^{T-1} |t\rangle |\psi_t\rangle$ . In the following, when we refer to the Turing machines encoded in the Hamiltonian ‘running’ on some input, we mean that the evolution produced by running the Turing machine on that input appears in the ground state as the corresponding computational history state.

If there are no other constraints, writing down a Hamiltonian whose ground state is the computational history state is straightforward. However, constructing such a Hamiltonian out of the local interactions of a many-body system is more involved. The construction method of ref. 15 was later substantially developed<sup>16</sup> and, after a long sequence of results<sup>17–19</sup>, culminated in the construction for 1D spin chains with translationally invariant, nearest-neighbour interactions presented in ref. 20.

For any quantum Turing machine<sup>21</sup> (QTM), an interaction  $h$  between neighbouring particles may be constructed<sup>20</sup> such that the ground state of the 1D translationally invariant Hamiltonian  $H_{GI} = \sum_{i=1}^N h_{i,i+1}$  is of the form  $\frac{1}{\sqrt{T}} \sum_{t=0}^{T-1} |t\rangle |\psi_t\rangle$ , where the ‘clock’ part of the computational history state  $|t\rangle \approx |1\rangle^{\otimes t} |0\rangle^{\otimes N-t}$  counts time in unary, and  $|\psi_t\rangle$  represents the state of the QTM after  $t$  time-steps. Moreover, the ground state energy may be taken equal to zero.

The translationally invariant Hamiltonians we are considering are completely specified by the finite number of matrix elements in the local interactions  $h_{\text{rows}}$ ,  $h_{\text{col}}$  and  $h_1$ . To encode the halting problem in

the Hamiltonian, we use quantum phase estimation<sup>22</sup> to encode any of the countably infinite possible inputs to the universal Turing machine (UTM) into these matrix elements.

### Quantum phase estimation

Given a unitary matrix  $U$ , the quantum phase estimation algorithm estimates an eigenvalue  $e^{2\pi i\varphi}$  of  $U$  to a given number of bits of precision (which must be chosen in advance). It is well-known<sup>22</sup> that if the number of bits of precision in the quantum phase estimation algorithm is greater than or equal to the number of digits in the binary fraction expansion of  $\varphi$ , then the quantum phase estimation algorithm, rather than estimating the phase approximately, will output all the digits of  $\varphi$  (written as a binary fraction) exactly.

We use this property to construct a family of QTMs  $P_n$ , indexed by  $n \in \mathbb{N}$ , with the following properties: (i) the number of internal states and tape symbols of  $P_n$  are independent of  $n$ ; and (ii) given a number  $N = 2^x - 1 \geq n$ , with  $x \in \mathbb{N}$ , as input (written in binary),  $P_n$  writes the binary expansion of  $n$  on its tape and then halts deterministically. (The reason for having the input  $N$  of this form will become clear later.) To construct  $P_n$ , we construct a QTM that uses the input  $N$  to determine how many digits of precision to use, then runs the quantum phase estimation algorithm on the single-qubit gate  $U = \begin{pmatrix} 1 & 0 \\ 0 & e^{2\pi i\varphi} \end{pmatrix}$ . The phase

$\varphi$  in  $U$  is determined by the transition rules of the QTM<sup>21</sup>. Choosing  $\varphi$  to be the rational number whose binary fraction expansion contains the digits of  $n$  (expressed in binary) achieves the desired behaviour for  $P_n$ . By ‘dovetailing’  $P_n$  with a UTM (that is, running  $P_n$  first, then running the UTM), the UTM runs on the input specified by  $\varphi$ .

The quantum computation carried out by  $P_n$  followed by the UTM is encoded in the Hamiltonian using the history state construction described above. The phase  $e^{2\pi i\varphi}$  being estimated then becomes one of the matrix elements of the Hamiltonian. The same happens with the  $e^{i\pi 2^{-|\varphi|}}$  term that appears in the inverse quantum Fourier transform—the key ingredient of the quantum phase estimation algorithm.

Finally, we must ensure that the  $|\psi_0\rangle$  component of the history state is correctly initialized to input of the form  $N = 2^x - 1$  (written in binary) required by  $P_n$ . But  $N = 2^x - 1$  in binary is simply a string of  $N$  ‘1’s, and it is easy to ensure that  $|\psi_0\rangle$  is the state  $|1\rangle^{\otimes N}$  using translationally invariant local interactions.

If we add an on-site interaction  $h_1 = |T\rangle\langle T|$  to the history-state Hamiltonian constructed above, which gives additional energy to the halting state  $|T\rangle$ , then its ground state will pick up additional energy if and only if the UTM halts. However, the ground state energy still converges to zero as  $L \rightarrow \infty$  in both cases (see Supplementary Information). The energy density therefore tends to zero in the thermodynamic limit, whether or not the UTM halts.

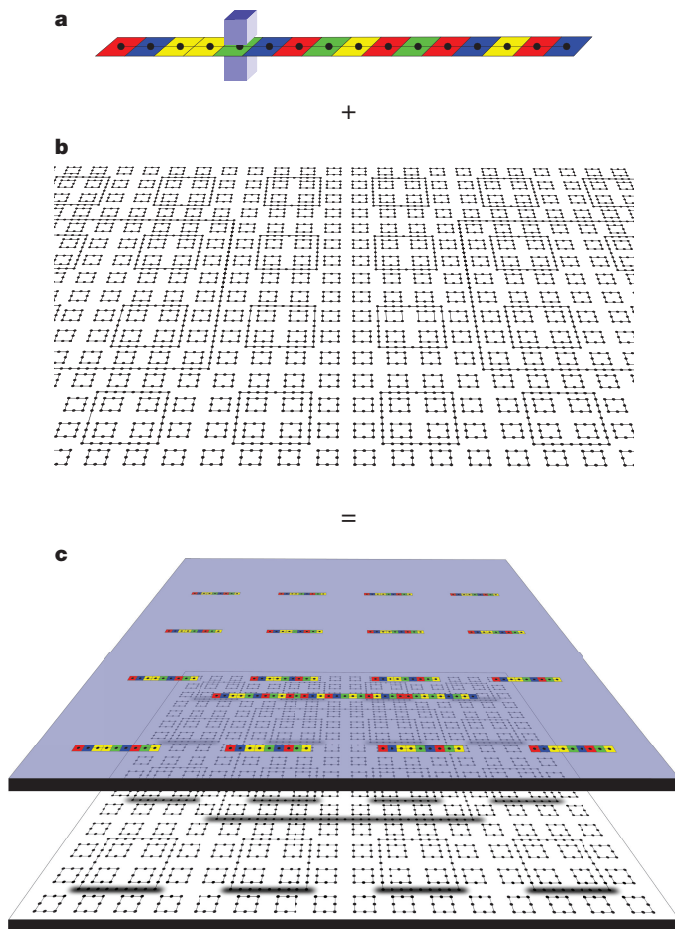
To remedy this, and amplify the difference between the halting and non-halting cases, we use the second spatial dimension and exploit Wang tilings.

### Quasi-periodic tilings

A Wang tile<sup>23</sup> is a square with markings along each edge. A tiling is then an arrangement of such tiles covering the whole plane, so that the markings on adjacent edges match. A tiling can easily be encoded in a ground state of a classical Hamiltonian on a 2D square lattice: by representing tile types by an orthogonal basis  $\{|T\rangle\}$  for the local Hilbert space  $\mathcal{H}_c$ , and choosing local interaction terms  $|T_i\rangle\langle T_i| \otimes |T_j\rangle\langle T_j|$  to give an energy penalty to all adjacent non-matching pairs of tiles  $T_i, T_j$ , a tiling of the plane is equivalent to a ground state with zero energy.

We prove, and subsequently exploit, very particular properties of the aperiodic Robinson tiling<sup>24</sup>, and combine them with the history-state Hamiltonian. Although the pattern of tiles in the Robinson tiling extends infinitely in all directions, it never repeats. More precisely, it contains periodically repeating subpatterns that form squares with sizes given by  $4^k$  for all  $k \in \mathbb{N}$  (see Fig. 3). This periodicity allows us to encode





**Figure 3 | Complete Hamiltonian construction.** a–c, The Robinson tiles enforce a recursive pattern of interlocking squares, the sizes of which are given by  $4^k$  for all  $k \in \mathbb{N}$  (b). As with any Wang tiling, we can readily represent this tiling as a classical Hamiltonian whose ground state has the same quasi-periodic structure. Because the set of tiles is fixed, the local dimension of this Hamiltonian is constant. By adding a ‘quantum layer’ on top of the Robinson-tiling Hamiltonian and choosing a suitable translationally invariant coupling between the layers, we effectively place copies of the QTM encoded in a 1D history-state Hamiltonian (a) along one edge of all of the squares. The ground state of this Hamiltonian consists of the Robinson tiling configuration in the tiling layer, with computational history states in the quantum layer along one edge of each square in the tiling (c). Each of these encodes the evolution of the same quantum phase estimation algorithm and UTM. The effective tape length available for each QTM is determined by the size of the square it ‘runs’ on.

in the ground state many copies of the UTM running on the same input  $\varphi$ , with tapes of all possible finite lengths and for every possible finite run time (see Fig. 3).

This encoding is achieved by sandwiching the 1D quantum history-state Hamiltonian  $h_q$  ‘on top of’ the Robinson-tiling Hamiltonian  $h_c$  to form two ‘layers’, so that the local Hilbert space at each site is  $\mathcal{H} = \mathcal{H}_c \otimes (\mathcal{H}_e \oplus \mathcal{H}_q)$  (where  $\mathcal{H}_e = |0\rangle$  is an additional energy level). One can then construct a Hamiltonian (see Supplementary Information) whose ground state is of the form  $|T\rangle_c \otimes |\psi\rangle_{eq}$ , where  $|T\rangle_c$  is a product state representing a classical configuration of the tiling layer and  $|\psi\rangle_{eq}$  contains—in a tensor product structure—computational history states along one edge (a ‘segment’) of all of the squares appearing in the configuration given by  $T$ . These computational history states are essentially the only constituents of  $|\psi\rangle_{eq}$  that contribute to the energy. The Hamiltonian also has an on-site interaction  $h_1 = |T\rangle\langle T|$  that gives an additional energy to the halting state of the Turing machine  $|T\rangle$ . Hence the ground state will pick up additional energy from all encoded Turing machines that halt. This energy still decreases as the relevant

system size increases, which, however, is now the size of the corresponding segment in the Robinson tiling (see Fig. 3), not the overall system size.

We now consider the ground state energy. If the UTM does not halt on input  $n$ , then  $|T\rangle_c$  is a valid tiling and for all segments that are larger than  $|n|$ , the ground state energy contribution is zero. The contribution for each segment smaller than  $|n|$  is some algebraic computable number. If  $\alpha(n)$  is the sum of the contributions of all segments smaller than  $|n|$ , then the addition of the constant energy shift  $h_1 = -\alpha(n)\mathbf{1}$  to the Hamiltonian makes the ground state energy density negative (but tending to zero from below as  $L \rightarrow \infty$ ) in the non-halting case (see Supplementary Information).

In the halting case, one of two things may happen. If  $|T\rangle_c$  is a valid tiling, then the number of squares large enough for the encoded Turing machine to halt grows quadratically with system size, and each square contributes a small but non-zero energy. Because such a state also picks up the energy contribution from segments of size smaller than  $|n|$ , the energy diverges with lattice size even after adding  $h_1 = -\alpha(n)\mathbf{1}$ . Hence the ground state energy density is strictly positive in the halting case, as desired.

Alternatively, one could try to reduce the energy by introducing defects in the tiling, which effectively ‘break’ some of the Turing machines so that they do not halt. However, we prove that the Robinson tiling is robust to such defects: a tile mismatch only affects the pattern of squares in a finite region around the defect, and each defect contributes  $\mathcal{O}(1)$  energy. We can choose the parameters (see Supplementary Information) to guarantee that introducing defects is energetically unfavourable. This completes the argument establishing our main result, Theorem 1.

Additional technical details can be found in the Supplementary Information.

## Discussion

We now discuss both the implications and the limitations of these results. This result is relevant to mathematical models of quantum many-body systems, as well as the behaviour of, and methods for treating, the thermodynamic limit. Moreover, it can also be seen as an indication of new physical phenomena.

An immediate consequence of the undecidability of the spectral gap is that there cannot exist an algorithm or a computable criterion that solves the spectral gap problem in general. Although algorithmic undecidability always concerns infinite families of systems, the axiomatic interpretation of the result also allows us to apply it to individual systems: there are particular Hamiltonians within these families for which one can neither prove nor disprove the presence of a gap, or of any other undecidable property. Unfortunately, our methods cannot pinpoint these particular cases, let alone prove that one of the aforementioned long-standing open problems is axiomatically undecidable.

A further consequence concerns the behaviour of the thermodynamic limit. In practice, we usually probe the idealized infinite thermodynamic limit by studying how the system behaves as we consider finite systems of increasing size. One often assumes that the systems, although finite, are so large that the asymptotic behaviour is already observed. In numerical simulations of condensed matter systems, one typically simulates finite systems of increasing size and extrapolates the asymptotic behaviour from the finite-size scaling<sup>25</sup>. Similarly, lattice quantum chromodynamics calculations simulate finite lattice spacings, and extrapolate the results to the continuum<sup>11</sup>. Renormalization group techniques accomplish something similar mathematically<sup>26</sup>; however, the undecidable quantum many-body models constructed in this work exhibit behaviour that defeats such approaches, in the following way. As the system size increases, the Hamiltonian will initially look like a gapless system, with the low-energy spectrum appearing to converge to a continuum. But at some threshold lattice size, a spectral gap of magnitude one will suddenly appear (or, vice versa, a gap will suddenly close<sup>27</sup>). Not only can the lattice size at which the system switches from

gapless to gapped be arbitrarily large, the threshold at which this transition occurs is uncomputable. The analogous implication also holds for all other undecidable low-temperature properties. Thus, any method of extrapolating the asymptotic behaviour from finite system sizes must fail in general.

This conclusion leads us directly to new physical phenomena. First, it hints at a new type of ‘phase transition,’ which is not driven by temperature or extrinsic local parameters, but by the size of the system. Some of the models constructed in the proof of Theorem 1 exhibit a drastic and abrupt change of properties when their size is increased beyond a certain scale. The scale at which this happens can be very large, and is not generally computable from the local description of the system. Second, our results show that certain quantum many-body models exhibit a radical form of instability. An arbitrarily small change in the parameters can cause the system to cross an arbitrary number of gapped/gapless transitions. In a sense, this phenomenon is the source of the undecidability in our models.

We finish with a closer look at some of the limitations of our results. First, all our results concern 2D (or higher-dimensional) systems. Although the majority of our construction is already 1D, we do not currently know whether the entire result holds in 1D as well. Second, although a theoretical model of a quantum many-body system is always an idealisation of the real physics, the models we construct in the proof of Theorem 1 are highly artificial. Whether the results can be extended to more natural models is yet to be determined. A related point is that we prove undecidability of the spectral gap (and other low-temperature properties) for Hamiltonians with a very particular form. We do not know how stable the results are to small deviations from this. This is a general issue with most many-body models; stability in this sense is not understood even for much simpler models such as the Ising model. Recent stability proofs only apply to certain types of frustration-free Hamiltonians<sup>28,29</sup>. Our results restrict the extent to which such stability results can be generalized. Similarly, we do not know whether the results hold for systems with low-dimensional local Hilbert spaces. Although the dimension  $d$  in Theorem 1 is fixed and finite, providing an estimate for it would be cumbersome and certainly involve large exponentials. However, the steps in the proof described above are not tailored to minimizing this dimension. Whether there is a non-trivial bound on the dimension of the local Hilbert space below which the spectral gap problem becomes decidable is an intriguing open question.

**Received 25 March; accepted 21 September 2015.**

- Schultz, T. D., Mattis, D. C. & Lieb, E. H. Two-dimensional Ising model as a soluble problem of many fermions. *Rev. Mod. Phys.* **36**, 856–871 (1964).
- Hastings, M. B. Lieb-Schultz-Mattis in higher dimensions. *Phys. Rev. B* **69**, 104431 (2004).
- Affleck, I., Kennedy, T., Lieb, E. H. & Tasaki, H. Valence bond ground states in isotropic quantum antiferromagnets. *Commun. Math. Phys.* **115**, 477–528 (1988).
- Haldane, F. D. M. Nonlinear field theory of large-spin Heisenberg antiferromagnets: semiclassically quantized solitons of the one-dimensional easy-axis Neel state. *Phys. Rev. Lett.* **50**, 1153–1156 (1983).
- Golinelli, O., Jolicoeur, Th. & Lacaze, R. Finite-lattice extrapolations for a Haldane-gap antiferromagnet. *Phys. Rev. B* **50**, 3037–3044 (1994).
- Anderson, P. W. Resonating valence bonds: a new kind of insulator? *Mater. Res. Bull.* **8**, 153–160 (1973).

- Yan, S., Huse, D. A. & White, S. R. Spin-liquid ground state of the  $S = 1/2$  kagome Heisenberg antiferromagnet. *Science* **332**, 1173–1176 (2011).
- Balents, L. Spin liquids in frustrated magnets. *Nature* **464**, 199–208 (2010).
- Han, T.-H. *et al.* Fractionalized excitations in the spin-liquid state of a kagome-lattice antiferromagnet. *Nature* **492**, 406–410 (2012).
- Jaffe, A. & Witten, E. in *The Millennium Prize Problems* (eds Carlson, J. A., Jaffe, A. & Wiles, A.) 129–152, <http://www.claymath.org/sites/default/files/yangmills.pdf> (Clay Mathematics Institute, American Mathematical Society, 2006).
- Rebbi, C. *Lattice Gauge Theories and Monte Carlo Simulations* (World scientific, 1983).
- Turing, A. M. On computable numbers, with an application to the Entscheidungsproblem. *Proc. Lond. Math. Soc.* **42**, 230–265 (1936).
- Gödel, K. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatsh. Math. Phys.* **38**, 173–198 (1931).
- Poonen, B. in *Interpreting Gödel: Critical Essays* (ed. Kennedy, J.) 211–241 (Cambridge Univ. Press, 2014).
- Feynman, R. P. Quantum mechanical computers. *Optics News* **11**, 11–20 (1985).
- Kitaev, A.Y., Shen, A.H. & Vyalii M.N. *Classical and Quantum Computation* Vol. 47 of *Graduate Studies in Mathematics* Ch. 14.4 (American Mathematical Society, 2002).
- Kempe, J., Kitaev, A. & Regev, O. The complexity of the local Hamiltonian problem. *SIAM J. Comput.* **35**, 1070–1097 (2006).
- Oliveira, R. & Terhal, B. M. The complexity of quantum spin systems on a two-dimensional square lattice. *Quantum Inf. Comput.* **8**, 900–924 (2008).
- Aharonov, D., Gottesman, D., Irani, S. & Kempe, J. The power of quantum systems on a line. *Commun. Math. Phys.* **287**, 41–65 (2009).
- Gottesman, D. & Irani, S. The quantum and classical complexity of translationally invariant tiling and Hamiltonian problems. In *Proc. 50th Annu. Symp. Found. Comput. Sci.* 95–104 (IEEE, 2009).
- Bernstein, E. & Vazirani, U. Quantum complexity theory. *SIAM J. Comput.* **26**, 1411–1473 (1997).
- Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* Ch. 5.2 (Cambridge Univ. Press, 2000).
- Wang, H. Proving theorems by pattern recognition. *Bell Syst. Tech. J.* **40**, 1–41 (1961).
- Robinson, R. M. Undecidability and nonperiodicity for tilings of the plane. *Invent. Math.* **12**, 177–209 (1971).
- Cardy, J. (ed.) *Finite-size Scaling* (Elsevier, 2012).
- Cardy, J. *Scaling and Renormalization in Statistical Physics* (Cambridge Univ. Press, 1996).
- Cubitt, T. S., Perez-Garcia, D. & Wolf, M. M. Undecidability of the spectral gap. Preprint at <http://arXiv.org/abs/1502.04573> (2015).
- Bravyi, S., Hastings, M. & Michalakis, S. Topological quantum order: stability under local perturbations. *J. Math. Phys.* **51**, 093512 (2010).
- Michalakis, S. & Zwolak, J. P. Stability of frustration-free Hamiltonians. *Commun. Math. Phys.* **322**, 277–302 (2013).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** T.S.C. thanks IBM. T. J. Watson Laboratory for their hospitality, and C. Bennett in particular for discussions about this work. T.S.C., D.P.-G. and M.M.W. thank the Isaac Newton Institute for Mathematical Sciences, Cambridge for their hospitality during the programme “Mathematical Challenges in Quantum Information”, where part of this work was carried out. T.S.C. is supported by the Royal Society. D.P.G. acknowledges support from MINECO (grant MTM2011-26912 and PRI-PIMCHI-2011-1071), Comunidad de Madrid (grant QITEMAD+-CM, ref. S2013/ICE-2801) and the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement no. 648913). This work was made possible through the support of grant no. 48322 from the John Templeton Foundation. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

**Author Contributions** All authors contributed extensively to the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.S.C. (t.cubitt@ucl.ac.uk).



# Signal integration by $\text{Ca}^{2+}$ regulates intestinal stem-cell activity

Hansong Deng<sup>1</sup>, Akos A. Gerencser<sup>1</sup> & Heinrich Jasper<sup>1</sup>

**Somatic stem cells maintain tissue homeostasis by dynamically adjusting proliferation and differentiation in response to stress and metabolic cues. Here we identify  $\text{Ca}^{2+}$  signalling as a central regulator of intestinal stem cell (ISC) activity in *Drosophila*. We show that dietary L-glutamate stimulates ISC division and gut growth. The metabotropic glutamate receptor (mGluR) is required in ISCs for this response, and for an associated modulation of cytosolic  $\text{Ca}^{2+}$  oscillations that results in sustained high cytosolic  $\text{Ca}^{2+}$  concentrations. High cytosolic  $\text{Ca}^{2+}$  concentrations induce ISC proliferation by regulating Calcineurin and CREB-regulated transcriptional co-activator (Crtc). In response to a wide range of dietary and stress stimuli, ISCs reversibly transition between  $\text{Ca}^{2+}$  oscillation states that represent poised or activated modes of proliferation, respectively. We propose that the dynamic regulation of intracellular  $\text{Ca}^{2+}$  levels allows effective integration of diverse mitogenic signals in ISCs to adapt their proliferative activity to the needs of the tissue.**

Somatic stem cells can shift between quiescent and active states, and between asymmetric and symmetric division modes according to the regenerative need of a tissue<sup>1–4</sup>. How stem cells decode and integrate various environmental inputs to adjust their proliferative behaviours remains an unresolved question with implications for our understanding of degenerative and neoplastic diseases. ISCs represent most mitosis-competent cells in the *Drosophila* intestinal epithelium, and can undergo symmetric or asymmetric divisions in response to dietary or stress stimuli, respectively<sup>1–3,5</sup>. Complex cell-autonomous and non-autonomous interactions between conserved signalling pathways (including Jak/Stat, Egfr, insulin receptor (InR) and JNK signalling) govern these responses<sup>1–3,5</sup>. Through asymmetric divisions, ISCs form lineage-restricted diploid progenitor cells known as enteroblasts, which differentiate into large polyploid enterocytes or small diploid enteroendocrine cells<sup>1,3,4</sup>. Symmetric divisions can be induced by insulin signalling and allow adaptive resizing of the gut after feeding<sup>1</sup>. Growth and physiology of flies is influenced by the protein concentration in the food, yet the role of specific nutrients in adaptive resizing has remained unexamined<sup>6,7</sup>. L-glutamate (L-Glu) is among the most abundant amino acids in proteins, and is a critical energy source for the intestine<sup>8</sup>. At the same time, it serves as a signalling molecule, stimulating specific membrane receptors in a wide range of cells<sup>9</sup>. L-Glu promotes cell proliferation in the intestinal epithelium of mice, mediated by the metabotropic glutamate receptor mGluR5 (ref. 10), yet it remains unclear whether and how this signal affects the proliferative activity of ISCs.

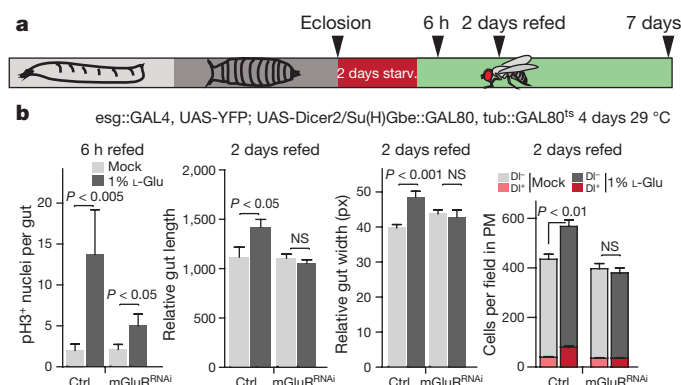
## Glutamate regulates gut growth through mGluR

To assess whether dietary L-Glu influences adaptive resizing of the intestinal epithelium, we used a modified version of the feeding protocol developed previously<sup>1</sup> (Fig. 1a and Extended Data Fig. 2a). As early as 4 h after starting feeding, flies maintained on 0.1% yeast supplemented with 1% L-Glu exhibited a higher mitotic index along the intestinal epithelium (both anterior and posterior gut) than controls, and after 2 days the posterior midgut significantly increased in length, width and cell density (Fig. 1b and Extended Data Fig. 2a, d; feeding rates were similar between the different food recipes; Extended Data Fig. 2b). L-Glu had to be ingested with the diet for these effects, as injection of 1% L-Glu into the animal did not increase ISC proliferation rates (Extended Data Fig. 2c). Food supplemented with other

amino acids, or in which the caloric content was increased using sugar, did not stimulate ISC proliferation (Extended Data Fig. 2e). L-Glu feeding also increased the growth rate of ISC lineages marked by Flp-out or MARCM lineage tracing systems<sup>11,12</sup> (Extended Data Fig. 2f–h).

We tested whether this L-Glu response is mediated by ionotropic receptors (NMDA (N-methyl-D-aspartate) and GluRIIA, B or C) or the one metabotropic L-Glu receptor (mGluR) encoded by the fly genome, and found that ISC-specific knockdown of mGluR or homozygosity for the mGluR loss-of-function allele *mGluR<sup>112b</sup>* impaired the response (Fig. 1b and Extended Data Fig. 2i). Desensitization or internalization of mGluR, or oxidative glutamate toxicity, may explain the observation that higher levels (5 or 10%) of L-Glu impaired ISC proliferation rather than activated it<sup>13,14</sup> (Extended Data Fig. 2a).

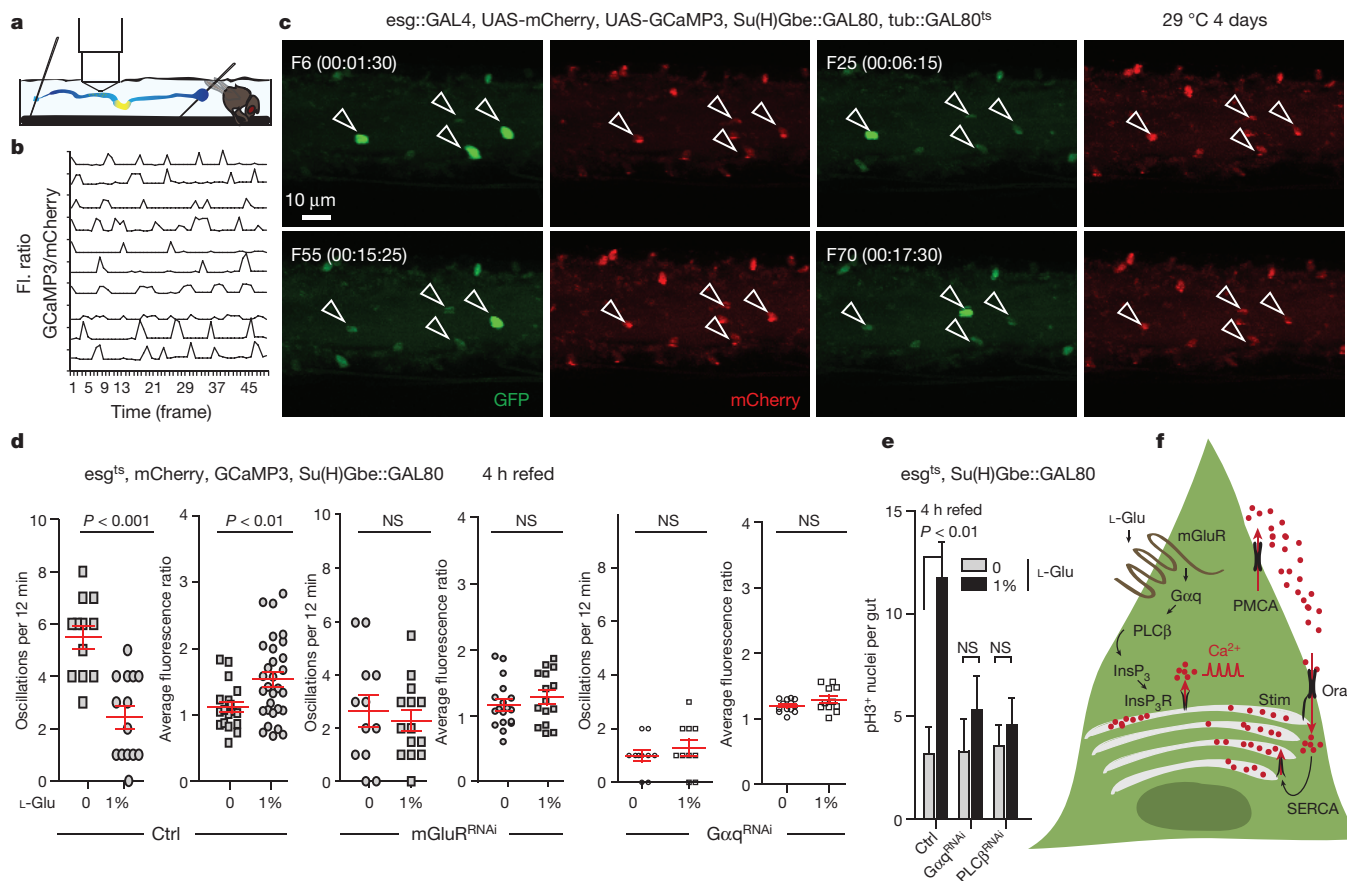
The concentration of freely available L-Glu in the intestinal epithelium is expected to reflect a balance between L-Glu absorption by



**Figure 1 | Glutamate regulates ISC proliferation and gut growth through mGluR.** **a**, Schematic of starvation/refeeding experiments performed.

**b**, Mitotic figures (phospho-histone H3 (pH3)-expressing nuclei), gut length, gut width, and number of DL-expressing and non-expressing cells per field in the posterior midgut (PM). Px, pixels. Data are mean and s.e.m. *P* values from Student's *t*-test; NS, not significant. For mitotic figures, *n* = 14 flies for each genotype, representative of 3 independent experiments shown. For gut length, width and cell numbers, *n* = 9 for each control (ctrl) condition, *n* = 12 flies for each *mGluR<sup>RNAi</sup>* condition, results are representative of 2 independent experiments.

<sup>1</sup>The Buck Institute for Research on Aging, 8001 Redwood Boulevard, Novato, California 94945, USA.



**Figure 2 | Dietary L-Glu influences ISC  $[\text{Ca}^{2+}]$  oscillations and proliferation through the mGluR/ $\text{G}\alpha\text{q}$ / $\text{PLC}\beta$  pathway.** **a**, Experimental setup for live recordings of  $\text{Ca}^{2+}$  oscillations in ISCs. **b**, Representative traces of GCaMP3/mCherry ratios in individual wild-type ISCs (pooled from 4 guts; 50 frames, 15 s intervals). FL, fluorescence. **c**, Representative frames from live recording (frame and time in minutes indicated). Arrowheads mark typical ISCs. GFP, green fluorescent protein. **d**,  $[\text{Ca}^{2+}]$  oscillation frequency and average GCaMP3/mCherry ratio in ISCs of flies of the indicated genotypes refed for 4 h. **e**, L-Glu-induced proliferation analysed after  $\text{G}\alpha\text{q}$  and  $\text{PLC}\beta$  were knocked down (using  $\text{G}\alpha\text{q}^{\text{RNAi}}$  and  $\text{PLC}\beta^{\text{RNAi}}$ , respectively) in ISCs for 7 days. Data are mean and s.e.m.  $P$  values from Student's  $t$ -test. In **d**, individual ISCs pooled from 4–5 guts;  $n = 12, 14, 18, 30, 12, 13, 17, 15, 9, 10, 13$  and 11 cells from left to right. Results are representative of 6 (**b, c**) and 3 (**d, e**) experiments. **f**, Regulation of  $\text{Ca}^{2+}$  homeostasis in non-excitable cells. GPCR/InsP<sub>3</sub>R cascade causes release of  $\text{Ca}^{2+}$  (red dots) from the ER. A decline of  $[\text{Ca}^{2+}]$  in the ER is sensed by the Stim–Orai complex, and induces extracellular  $\text{Ca}^{2+}$  influx into the cytoplasm (SOCE). Excessive cytoplasmic  $\text{Ca}^{2+}$  is pumped into the ER by SERCA or out of the cell by PMCA.

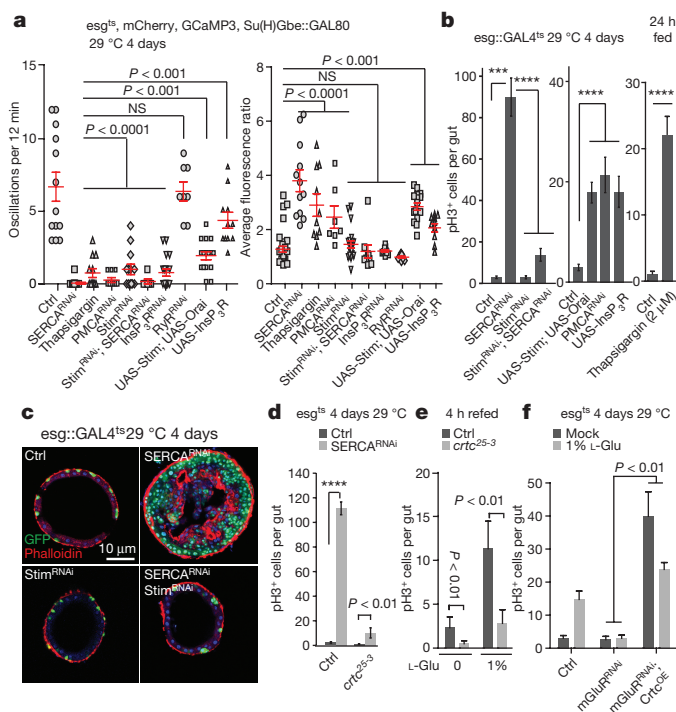
enterocytes and input from the diet. In the nervous system, excess L-Glu is recycled by astrocytes through Excitatory amino acid transporters (Eaat proteins)<sup>15</sup>. In mammals, EAAT3 (also known as EAAC1 and SLC1A1) is expressed in the apical, brush border membrane of enterocytes throughout the small intestine<sup>16</sup>. RNA sequencing (RNA-seq), quantitative reverse transcription PCR (qRT-PCR), *in situ* hybridization and an Eaat1::GAL4 reporter<sup>17</sup> suggested expression of *Eaat1* in enterocytes of the fly gut, with specifically increased activity in the anterior midgut (Extended Data Fig. 3a–c and not shown). Silencing *Eaat1* in enterocytes using the enterocyte-specific driver NP1::GAL4 (ref. 18) increased mitotic figures and the number of cells expressing the ISC marker Delta throughout the gut without triggering the canonical Jak/Stat-mediated enterocyte stress response or inducing apoptosis in enterocytes (Extended Data Fig. 3d, f, g). Overexpression of *Eaat1* significantly reduced the ISC response to L-Glu feeding (Extended Data Fig. 3e), suggesting that L-Glu resorption through Eaat1 influences free L-Glu levels in the intestine, modulating ISC proliferation and adaptive resizing. The limited glutamatergic innervation of the midgut epithelium suggests that direct regulation of ISCs by glutamatergic neurons is unlikely (Extended Data Fig. 4a–d).

### $\text{Ca}^{2+}$ oscillations in ISCs

mGluR, a G-protein-coupled receptor (GPCR), influences cytosolic  $\text{Ca}^{2+}$  levels via phospholipase C (PLC) and inositol-1,4,5-trisphosphate

(InsP<sub>3</sub>)<sup>19</sup>. Cytosolic  $[\text{Ca}^{2+}]$  is low owing to active transport that sequesters  $\text{Ca}^{2+}$  into the endoplasmic reticulum (ER), the mitochondria or the extracellular space. Stimulation by growth factors, hormones or neurotransmitters triggers  $\text{Ca}^{2+}$  release from these stores, leading to acute spikes in cytosolic  $[\text{Ca}^{2+}]$  as well as  $[\text{Ca}^{2+}]$  oscillations that influence a wide range of cellular functions<sup>20–23</sup>. We used UAS-GCaMP3, a genetically encoded  $\text{Ca}^{2+}$  reporter<sup>24,25</sup>, to test whether intracellular  $[\text{Ca}^{2+}]$  in ISCs responds to dietary L-Glu and to mGluR activity. Variable activity of this reporter in ISCs (Extended Data Fig. 4e–g) suggested fluctuations of  $[\text{Ca}^{2+}]$  that we confirmed by live imaging using two-photon microscopy (Fig. 2a–c). Quantification of  $[\text{Ca}^{2+}]$  in individual ISCs (as the fluorescence emission ratio between GCaMP and co-expressed mCherry) revealed frequent and robust  $[\text{Ca}^{2+}]$  oscillations in ISCs of wild-type intestines (Supplementary Discussion, Fig. 2b, c, Extended Data Fig. 5a, e and Supplementary Videos 1 and 2). Experiments using  $\text{Ca}^{2+}$  channel inhibitors and  $\text{Ca}^{2+}$  chelators, as well as genetic knockdown of InsP<sub>3</sub> receptors, suggest that these  $[\text{Ca}^{2+}]$  oscillations depend on influx through plasma membrane  $\text{Ca}^{2+}$  channels and on  $\text{Ca}^{2+}$  release from intracellular stores through InsP<sub>3</sub> receptors (Supplementary Discussion, Fig. 3a and Extended Data Fig. 5b–d). We confirmed reliability of GCaMP3 by recording  $\text{Ca}^{2+}$  oscillations using a bi-cistronic tdTomato-V2-GCaMP5 construct<sup>26,27</sup>, and a new, ultra-sensitive (higher affinity, dynamic range and brightness)  $\text{Ca}^{2+}$  reporter, IVS-GCaMP6s<sup>28</sup> (Extended Data Fig. 5e and Supplementary Video 2).





**Figure 3 | CaN/CRTC regulates ISC proliferation in response to increased cytosolic  $\text{Ca}^{2+}$ .** **a**,  $\text{Ca}^{2+}$  oscillation patterns in ISCs perturbed as indicated. **b**, Mitotic figures in guts of 2–3-day-old flies maintained at 29 °C for 4 days. **c**, Cross section of posterior midguts of the indicated genotypes. Visceral muscle identified by phalloidin (red), and ISCs/enteroblasts by GFP (green). DAPI, blue. **d**, Mitotic figures in control or *crtc* null mutants (*crtc<sup>25-3</sup>*) combined with SERCA knockdown (*SERCA<sup>RNAi</sup>*). **e**, Mitotic figures in control or *crtc<sup>25-3</sup>* homozygotes after refeeding. **f**, Mitotic figures in indicated genotypes after refeeding. Data are mean and s.e.m. In **a**, individual ISCs pooled from 4–5 different guts;  $n = 13, 12, 13, 10, 16, 12, 14, 8, 17, 11, 21, 12, 11, 10, 16, 12, 8, 6, 20$  and 12 cells. In **b**,  $n = 22, 24, 15, 13, 16, 15, 22, 18, 12$  and 16 guts. In **d**,  $n = 12, 22, 11$  and 19 guts. In **e**,  $n = 12, 8$  and 8 guts. In **f**,  $n = 10, 12, 11, 11, 10$  and 12 guts (from left to right in all panels). \*\*\* $P < 0.001$ ; \*\*\*\* $P < 0.0001$ , one-way analysis of variance (ANOVA) (**a**, **b** (left and middle), **d**, **f**) and Student's *t*-test (**b** (right), **d**). Results are representative of 3 (**a**, **e**) and 2 (**b**, **d**, **f**) experiments.

Refeeding with 1% L-Glu-supplemented food reduced  $[\text{Ca}^{2+}]$  oscillation frequency and increased the average GCaMP3 signal intensity in wild-type ISCs (Fig. 2d, Extended Data Fig. 6a and Supplementary Videos 3 and 4). Knockdown of *mGluR* resulted in reduced oscillation frequency while maintaining low signal intensity under mock conditions, and prevented the L-Glu-induced increase in signal intensity (Fig. 2d, Extended Data Fig. 6a, c and Supplementary Videos 5 and 6). These data suggested that increased ISC proliferation is associated with increased cytosolic  $[\text{Ca}^{2+}]$ . Accordingly, both oscillations and signal intensity were reduced in feeding conditions in which ISC proliferation is inhibited (Extended Data Fig. 6b; compare with Extended Data Fig. 2a).

The *Drosophila* genome encodes six different  $\text{G}\alpha$  subunits, three different  $\text{G}\beta$  subunits and two  $\text{G}\gamma$  subunits.  $\text{PLC}\beta$ ,  $\text{G}\alpha_q$ ,  $\text{G}\beta 13F$  and  $\text{G}\gamma 1$  (but not  $\text{G}\alpha_o$ ,  $\text{G}\alpha_i$  and  $\text{G}\alpha_s$ ) are required for  $[\text{Ca}^{2+}]$  oscillations in ISCs (Extended Data Fig. 6d, e and Supplementary Video 7), and knockdown of  $\text{G}\alpha_q$  or  $\text{PLC}\beta$  inhibited L-Glu-induced changes in  $[\text{Ca}^{2+}]$  oscillations and ISC proliferation (Fig. 2d, e). A canonical GPCR signalling pathway through heterotrimeric G proteins and  $\text{PLC}\beta$  thus regulates cytosolic  $[\text{Ca}^{2+}]$  and proliferation of ISCs in response to dietary L-Glu (Fig. 2f).

### Increased cytosolic $[\text{Ca}^{2+}]$ triggers ISC proliferation

$\text{Ca}^{2+}$  concentrations in the ER are dynamically controlled by Store-operated calcium entry (SOCE)<sup>20–22</sup>. Cytosolic  $\text{Ca}^{2+}$  is pumped into

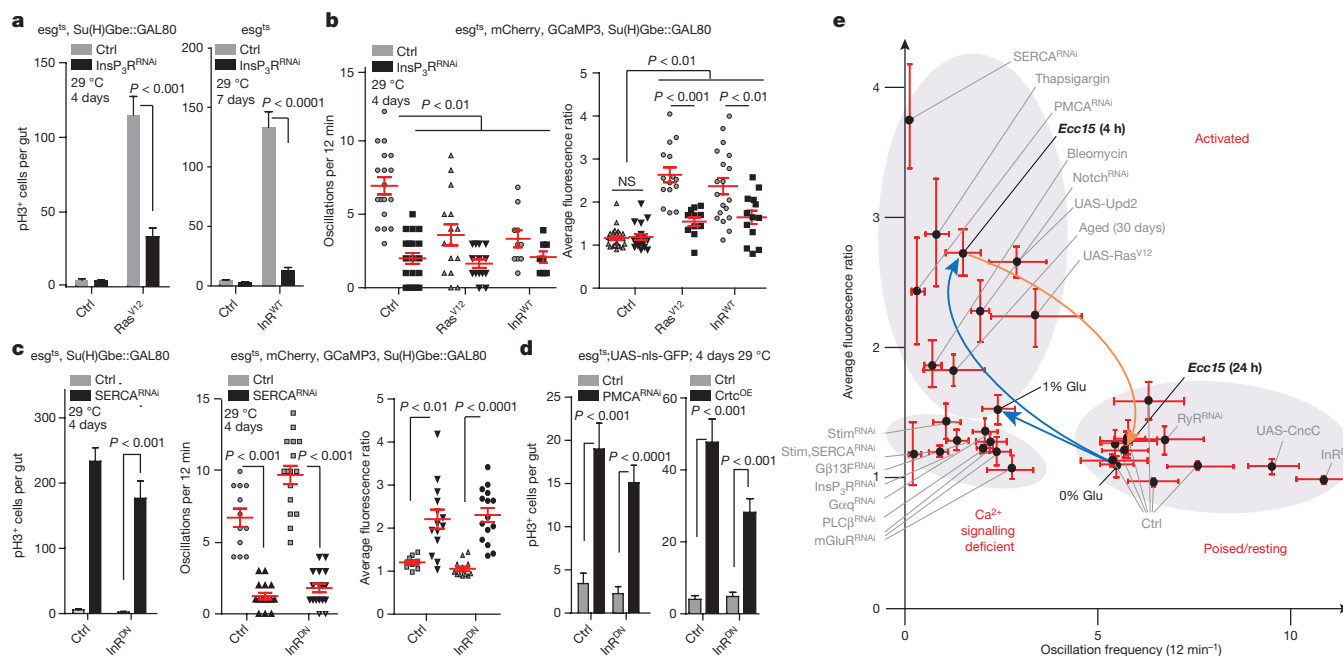
the ER by Sarco-endoplasmic reticulum calcium ATPase (SERCA), or out of the cell by Sodium calcium exchanger (Ncx) or Plasma membrane calcium ATPase (PMCA)<sup>20–22</sup>. A decrease in ER  $[\text{Ca}^{2+}]$  is sensed by Stromal interaction molecule (Stim), an ER membrane protein that opens the plasma membrane  $\text{Ca}^{2+}$  channel Orai, allowing influx of extracellular  $\text{Ca}^{2+}$  into the cytosol<sup>20–22</sup>. This coupled influx/efflux process is essential for  $\text{Ca}^{2+}$  oscillations in non-excitable cells<sup>20–22</sup> and is conserved in *Drosophila* (Fig. 2f and Extended Data Fig. 1b). Genetic or pharmacological perturbation of conserved SOCE components revealed that  $[\text{Ca}^{2+}]$  oscillations in ISCs are regulated by the canonical GPCR/InsP<sub>3</sub>/SOCE pathway, while receptors that influence membrane polarization do not contribute to  $\text{Ca}^{2+}$  signalling in these cells (Supplementary Discussion, Fig. 3a, Extended Data Figs 6f and 7a and Supplementary Videos 8 and 9). Perturbations in which oscillations were impaired while cytosolic  $[\text{Ca}^{2+}]$  increased (for example, knocking down SERCA or PMCA) strongly stimulated ISC proliferation (Fig. 3b, c and Extended Data Fig. 7b–d; see Supplementary Discussion for detailed description of genetic perturbations). This could be prevented by inhibiting the increase in cytosolic  $[\text{Ca}^{2+}]$  (by also knocking down Stim; Fig. 3b, c and Extended Data Fig. 7e, f). Sustained increased cytosolic  $[\text{Ca}^{2+}]$  is thus sufficient and required to induce ISC proliferation. ISC proliferation is also induced by overexpression of Stim and Orai or InsP<sub>3</sub>R, conditions that increase cytosolic  $[\text{Ca}^{2+}]$  independently of SERCA (Fig. 3a, b and Extended Data Fig. 7f).

SERCA deficiency can cause mis-folding and mis-processing of Notch receptors, resulting in ISC tumours<sup>29,30</sup>. We therefore assessed a possible role for Notch modulation in the proliferative response of ISCs to increased  $[\text{Ca}^{2+}]$ . Studies of Notch reporter expression, of Notch-dependent differentiation phenotypes, and of Notch gain-of-function conditions suggest that the acute mitogenic effect of sustained increased cytosolic  $[\text{Ca}^{2+}]$  can be decoupled from longer-term effects of SERCA deficiency on Notch signalling, and that increasing  $[\text{Ca}^{2+}]$  through SERCA-independent means or inducing ISC proliferation by overexpressing the  $\text{Ca}^{2+}$  effector Crtc (see below) does not influence Notch activity (Supplementary Discussion and Extended Data Fig. 8). SERCA deficiency further does not induce ISC proliferation by triggering an ER stress response (Supplementary Discussion and Extended Data Fig. 9a, b), and knockdown of SERCA or PMCA, or overexpression of Crtc, does not result in increased diphospho-ERK levels in ISCs (Extended Data Fig. 9c). Our experiments thus fail to detect deregulation of Notch, ERK and ER stress signalling as associated with the acute proliferative response of ISCs to increased cytosolic  $[\text{Ca}^{2+}]$ . Although an as-yet undetected perturbation of these pathways by short-term knockdown of SERCA or PMCA cannot be ruled out, we therefore sought to identify additional  $\text{Ca}^{2+}$ -sensitive regulators that may mediate the proliferative response of ISCs.

### CaN/CRTC regulate ISC proliferation in response to $[\text{Ca}^{2+}]$

Silencing the regulatory subunit *CanB2* or the catalytic subunit *CanA1* of the  $\text{Ca}^{2+}$ -dependent phosphatase Calcineurin (CaN), but not the  $\text{Ca}^{2+}$ -dependent kinases CaMKI and CaMKII, significantly abrogated ISC proliferation induced by either RNA interference (RNAi) targeting SERCA (*SERCA<sup>RNAi</sup>*), PMCA (*PMCA<sup>RNAi</sup>*), or by co-overexpression of Stim and Orai (*Stim<sup>OE</sup>* and *Orai<sup>OE</sup>*) (Extended Data Fig. 9d, e). Knockdown of other CaN subunits encoded by the *Drosophila* genome (*CanA-14F*, *Pp2B-14D* and *CanB*) did not limit ISC proliferation (not shown), but overexpression of any of the three constitutively activated catalytic subunits<sup>31</sup> was sufficient to promote ISC proliferation (Extended Data Fig. 9f). MARCM clones generated by ISCs homozygous for the *CanB2<sup>KO</sup>* null allele<sup>32</sup> or expressing *CanB2<sup>RNAi</sup>* grow slower than wild-type clones, while overexpressing constitutively active *CanA-14F* promotes clone growth (Extended Data Fig. 9g).

The transcription factor Crtc (CREB-regulated transcription co-activator, also known as transducer of regulated CREB activity,



**Figure 4** |  $\text{Ca}^{2+}$  signalling integrates stress and mitogenic signals to stimulate ISC proliferation. **a**, Mitotic figures in indicated genotypes. **b**, Frequency of  $[\text{Ca}^{2+}]$  oscillations and average fluorescence ratio in individual ISCs of indicated genotypes. **c**, Quantification of mitotic figures in the midgut, and of  $[\text{Ca}^{2+}]$  in ISCs of the indicated genotypes. **d**, Quantification of mitotic figures in the midgut of the indicated genotypes. **e**, ISC  $[\text{Ca}^{2+}]$  oscillation patterns segregate into three different modes (shaded areas) that correlate with proliferative activity. ISCs with perturbed  $\text{Ca}^{2+}$  signalling machinery are proliferation deficient. Blue arrows indicate ISC activation after L-Glu feeding or *Ecc15* infection. Orange arrow indicates

return to quiescence 24 h after *Ecc15* infection. Data are and s.e.m. *P* values by Student's *t*-test (**a**, **c**, **d**), and one-way ANOVA (**b**). In **a**,  $n = 12, 12, 11, 14, 20, 14, 14, 15$  guts. In **b**,  $n = 18, 27, 15, 15, 10, 11, 25, 20, 15, 22, 22$  and 17 cells. In **c**, for mitotic figures,  $n = 12, 18, 10$  and 9 guts. For  $\text{Ca}^{2+}$  recordings, individual ISCs pooled from 4–5 guts;  $n = 11, 12, 15, 19, 8, 14, 10$  and 16 cells. In **d**,  $n = 9, 12, 11, 11, 12, 12, 13$  and 12 guts (from left to right in all panels). Results representative of 2 (**a**, **c** (mitotic figures), **d**) and 3 (**b**, **c** ( $\text{Ca}^{2+}$  recordings)) experiments. In **e**, mean and s.e.m. of oscillation frequency and average fluorescence ratio are plotted for each condition.

Torc; not to be confused with Target of rapamycin complex I, TorC1)<sup>33</sup> is a conserved substrate of CaN, which promotes its nuclear translocation. Overproliferation of ISCs expressing *SERCA*<sup>RNAi</sup>, *Pp2B-14D*<sup>act</sup> or *PMCA*<sup>RNAi</sup> is suppressed in flies homozygous for the null allele *crtc*<sup>25-3</sup> (ref. 34; Fig. 3d and Extended Data Fig. 9h, j), while overexpression of wild-type *Crtc* or constitutively active *Crtc*<sup>35</sup> alone is sufficient to induce ISC proliferation (Extended Data Fig. 9i, j). Finally, L-Glu-induced ISC proliferation is not observed in *crtc*<sup>25-3</sup> homozygotes (Fig. 3e), while *Crtc* is sufficient to induce proliferation independently of *mGluR*, *Gqα* and *InsP3R* (also known as *Itp-r83A*) (Fig. 3f and Extended Data Fig. 9k). Overexpressing the conserved *Crtc* effector CREB, *CrebB-17A*, also causes ISC over-proliferation (Extended Data Fig. 9l; note that knockdown of CREB and its partner CBP leads to ISC loss, suggesting a broader role of these factors in ISC maintenance).

### $[\text{Ca}^{2+}]$ integrates mitogenic signals in ISCs

ISCs sample a variety of environmental, systemic and local cues to adjust proliferative activity to the needs of the tissue, and we asked whether  $[\text{Ca}^{2+}]$  modulation is a general feature of conditions that promote ISC proliferation. We monitored  $\text{Ca}^{2+}$  oscillations in the following mitogenic conditions<sup>1-3,5</sup>: overexpression of *InR*, *Ras*<sup>V12</sup> (a constitutively active allele of *Ras*) or *Upd2* (a ligand of the Jak/Stat pathway that stimulates ISC proliferation after tissue damage), infection with *Ecc15* (*Erwinia carotovora carotovora* 15; oral infection with *Ecc15* transiently stimulates ISC proliferation), loss of Notch (resulting in frequent symmetric divisions of ISCs), and bleomycin treatment (which causes DNA damage in enterocytes, resulting in compensatory proliferation of ISCs). In all cases, oscillation frequency decreased, while the average signal intensity increased (Extended Data Fig. 10a–c). Changes in  $\text{Ca}^{2+}$  oscillations are reversible: concomitant with the transient induction of proliferation

after *Ecc15* infection<sup>36</sup>, signal intensity increases and frequency decreases at 4 h after infection, returning to basal states at 24 h (when ISC proliferation subsides<sup>36</sup>; Extended Data Fig. 10c). Similar results were obtained with the bi-cistronic dTomato-V2-GCaMP5 construct (Extended Data Fig. 10d), and changes in oscillation pattern were also found in ISCs from old guts (30 days old; Extended Data Fig. 10a, b), which exhibit general hyperproliferation of ISCs<sup>2,5</sup>. Conditions in which ISC proliferation was inhibited (overexpression of *CncC*<sup>37</sup> or of dominant-negative *InR* (*InR*<sup>DN</sup>)<sup>38</sup>) increase oscillation frequency without changing average signal intensity (Extended Data Fig. 10a, b, e and Supplementary Video 10). Expression of *InR*<sup>DN</sup> (*InR*<sup>K1409A</sup>) in ISCs inhibits bleomycin or feeding-induced proliferation (Extended Data Fig. 10f).

To test whether  $\text{Ca}^{2+}$  signalling is involved in the integration of different upstream stimuli that control ISC proliferation, we examined whether it is required for proliferation induced by the overexpression of *InR* and *Ras*<sup>V12</sup> (representing nutrient- and growth-factor-induced ISC proliferation) or of the Jun-amino-terminal kinase kinase Hemipterous (*Hep*, representing stress-induced ISC proliferation), by knockdown of Notch, by *Ecc15* infection, or by exposure to bleomycin. While these perturbations strongly activated wild-type ISCs, activation of ISCs in *Orai*, *InsP3R*, *CanB2*, *Stim* or *crtc* loss-of-function conditions was significantly reduced (Fig. 4a, b and Extended Data Figs 8i and 10g, i, j). Accordingly, loss of *InsP3R* or *Orai* prevented the increase in cytosolic  $[\text{Ca}^{2+}]$  by these perturbations (Fig. 4b and Extended Data Fig. 10h). *mGluR* deficiency did not impair bleomycin- or *Ecc15*-induced proliferation and the associated cytosolic  $[\text{Ca}^{2+}]$  increase (Extended Data Fig. 10i, k).

Increasing cytosolic  $[\text{Ca}^{2+}]$  by knocking down *SERCA* or *PMCA*, or overexpressing *Crtc* is sufficient to bypass the requirement for *InR* in ISC proliferation (Fig. 4c, d), as well as the requirement for Jak/Stat signalling in *Ecc15*-induced ISC proliferation (Extended Data Fig. 10l).



Knockdown of SERCA was not sufficient to induce proliferation in loss-of-function conditions for the ERK-regulated transcription factor Fos<sup>39</sup>, however, demonstrating a parallel requirement for Crtc and RTK/ERK-regulated transcription factors in the proliferative response to mitogens (Extended Data Fig. 10m).

When [Ca<sup>2+</sup>] oscillation frequency is plotted relative to GCaMP3 fluorescence across a range of conditions, proliferative and resting states of ISCs segregate into two distinct modes (Fig. 4e and Extended Data Fig. 10n): quiescent ISCs poised for division (as in young guts under homeostatic conditions) show frequent and robust Ca<sup>2+</sup> oscillations with a low level of cytosolic Ca<sup>2+</sup>, while highly proliferative ISCs have reduced Ca<sup>2+</sup> oscillations with high levels of cytosolic Ca<sup>2+</sup>. Conversely, ISCs in which Ca<sup>2+</sup> signalling is impaired by knocking down components of the mGluR/PLC $\beta$ /InsP<sub>3</sub> pathway show a third pattern: reduced oscillation frequency with low [Ca<sup>2+</sup>], and are proliferation deficient. ISC modes segregate regardless of whether average fluorescence intensity (calculated based on raw fluorescence values, Fig. 4e) or average baseline (calculated based on Gaussian fits, Extended Data Fig. 10n) are plotted against oscillation frequencies, but not when mean local amplitudes are plotted (Extended Data Fig. 10n), indicating that the primary driver of ISC proliferation is not the intensity of individual Ca<sup>2+</sup> spikes, but the increase in average [Ca<sup>2+</sup>] in the ISC cytosol.

## Discussion

We propose that the dynamic control of cytosolic [Ca<sup>2+</sup>] is crucial for the ability of ISCs to undergo rapid and reversible activation and to control their proliferative activity dynamically in response to a wide range of nutrient and stress signals. InsP<sub>3</sub>/Ca<sup>2+</sup>/CaN/Crtc signalling coordinates and integrates cellular responses to these signals (Extended Data Fig. 1a). Exploring this integration further by assessing the role of Ca<sup>2+</sup> signalling in mediating proliferative responses to ER stress<sup>40</sup> and hemocyte-derived growth factors<sup>41</sup>, as well as in the shift from asymmetric to symmetric division modes in response to insulin-like peptides<sup>1</sup> will be of interest.

Our data suggest that cytosolic [Ca<sup>2+</sup>] changes in a wide range of conditions, and a detailed characterization of the kinetics of these changes, of the relative requirement for Crtc in the mitogenic response to Egf- and insulin-like ligands, as well as of the contribution of Notch signalling deregulation in [Ca<sup>2+</sup>]-induced ISC proliferation will shed further light onto the integration of the complex array of mitogenic signals sensed by ISCs.

In vertebrates, high plasma L-Glu levels and expression and activity of mGluR homologues are associated with malignancy<sup>42,43</sup>, and our results indicate that chronic mGluR stimulation in stem and progenitor-like cells may be a driver of proliferative deregulation. This may be particularly relevant in the intestine, where L-Glu is not only introduced by the diet, but can also be a product of the commensal microbiome<sup>44,45</sup>. Understanding mGluR signalling in ISCs is thus likely to provide significant new leads for possible therapies of intestinal cancers and inflammatory diseases.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 21 December 2014; accepted 30 October 2015.**

**Published online 2 December 2015.**

- O'Brien, L. E., Soliman, S. S., Li, X. & Bilder, D. Altered modes of stem cell division drive adaptive intestinal growth. *Cell* **147**, 603–614 (2011).
- Buchon, N., Broderick, N. A. & Lemaître, B. Gut homeostasis in a microbial world: insights from *Drosophila melanogaster*. *Nature Rev. Microbiol.* **11**, 615–626 (2013).
- Biteau, B., Hochmuth, C. E. & Jasper, H. Maintaining tissue homeostasis: dynamic control of somatic stem cell activity. *Cell Stem Cell* **9**, 402–411 (2011).
- Lemaître, B. & Miguel-Alíaga, I. The digestive tract of *Drosophila melanogaster*. *Annu. Rev. Genet.* **47**, 377–404 (2013).
- Ayyaz, A. & Jasper, H. Intestinal inflammation and stem cell homeostasis in aging *Drosophila melanogaster*. *Front. Cell. Infect. Microbiol.* **3**, 98 (2013).

- Skorupa, D. A., Dervisevendic, A., Zwiener, J. & Pletcher, S. D. Dietary composition specifies consumption, obesity, and lifespan in *Drosophila melanogaster*. *Aging Cell* **7**, 478–490 (2008).
- Lee, K. P. et al. Lifespan and reproduction in *Drosophila*: new insights from nutritional geometry. *Proc. Natl Acad. Sci. USA* **105**, 2498–2503 (2008).
- Reeds, P. J., Burrin, D. G., Stoll, B. & Jahoor, F. Intestinal glutamate metabolism. *J. Nutr.* **130**, 978S–982S (2000).
- Newsholme, P., Procopio, J., Lima, M. M., Pithon-Curi, T. C. & Curi, R. Glutamine and glutamate—their central role in cell metabolism and function. *Cell Biochem. Funct.* **21**, 1–9 (2003).
- Xiao, W. et al. Glutamate prevents intestinal atrophy via luminal nutrient sensing in a mouse model of total parenteral nutrition. *FASEB J.* **28**, 2073–2087 (2014).
- Jiang, H. et al. Cytokine/Jak/Stat signaling mediates regeneration and homeostasis in the *Drosophila* midgut. *Cell* **137**, 1343–1355 (2009).
- Lee, T. & Luo, L. Mosaic analysis with a repressible cell marker (MARCM) for *Drosophila* neural development. *Trends Neurosci.* **24**, 251–254 (2001).
- Lewerenz, J. et al. The cystine/glutamate antiporter system x<sub>c</sub><sup>−</sup> in health and disease: from molecular mechanisms to novel therapeutic opportunities. *Antioxid. Redox Signal.* **18**, 522–555 (2013).
- Freedman, N. J. & Lefkowitz, R. J. Desensitization of G protein-coupled receptors. *Recent Prog. Horm. Res.* **51**, 319–351; discussion 352–313 (1996).
- Danbolt, N. C. Glutamate uptake. *Prog. Neurobiol.* **65**, 1–105 (2001).
- Burrin, D. G. & Stoll, B. Metabolic fate and function of dietary glutamate in the gut. *Am. J. Clin. Nutr.* **90**, 850S–856S (2009).
- Rival, T. et al. Decreasing glutamate buffering capacity triggers oxidative stress and neuropil degeneration in the *Drosophila* brain. *Curr. Biol.* **14**, 599–605 (2004).
- Li, H., Qi, Y. & Jasper, H. Dpp signaling determines regional stem cell identity in the regenerating adult *Drosophila* gastrointestinal tract. *Cell Rep* **4**, 10–18 (2013).
- McBride, S. M. et al. Pharmacological rescue of synaptic plasticity, courtship behavior, and mushroom body defects in a *Drosophila* model of fragile X syndrome. *Neuron* **45**, 753–764 (2005).
- Brini, M. & Carafoli, E. Calcium pumps in health and disease. *Physiol. Rev.* **89**, 1341–1378 (2009).
- Dupont, G., Combettes, L., Bird, G. S. & Putney, J. W. Calcium oscillations. *Cold Spring Harb. Perspect. Biol.* **3**, a004226 (2011).
- Clapham, D. E. Calcium signaling. *Cell* **131**, 1047–1058 (2007).
- Shim, J. et al. Olfactory control of blood progenitor maintenance. *Cell* **155**, 1141–1153 (2013).
- Tian, L. et al. Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nature Methods* **6**, 875–881 (2009).
- Akerboom, J. et al. Optimization of a GCaMP calcium indicator for neural activity imaging. *J. Neurosci.* **32**, 13819–13840 (2012).
- Wong, C. O. et al. A TRPV channel in *Drosophila* motor neurons regulates presynaptic resting Ca<sup>2+</sup> levels, synapse growth, and synaptic transmission. *Neuron* **84**, 764–777 (2014).
- Daniels, R. W., Rossano, A. J., Macleod, G. T. & Ganetzky, B. Expression of multiple transgenes from a single construct using viral 2A peptides in *Drosophila*. *PLoS ONE* **9**, e100637 (2014).
- Chen, T. W. et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).
- Roti, G. et al. Complementary genomic screens identify SERCA as a therapeutic target in NOTCH1 mutated cancer. *Cancer Cell* **23**, 390–405 (2013).
- Periz, G. & Fortini, M. E. Ca<sup>2+</sup>-ATPase function is required for intracellular trafficking of the Notch receptor in *Drosophila*. *EMBO J.* **18**, 5983–5993 (1999).
- Nakai, Y. et al. Calcineurin and its regulator Sra/DSCR1 are essential for sleep in *Drosophila*. *J. Neurosci.* **31**, 12759–12766 (2011).
- Takeo, S., Tsuda, M., Akahori, S., Matsuo, T. & Aigaki, T. The calcineurin regulator sra plays an essential role in female meiosis in *Drosophila*. *Curr. Biol.* **16**, 1435–1440 (2006).
- Altarejos, J. Y. & Montminy, M. CREB and the CREB co-activators: sensors for hormonal and metabolic signals. *Nature Rev. Mol. Cell Biol.* **12**, 141–151 (2011).
- Wang, B. et al. The insulin-regulated CREB coactivator TORC promotes stress resistance in *Drosophila*. *Cell Metab.* **7**, 434–444 (2008).
- Hirano, Y. et al. Fasting launches CRTC to facilitate long-term memory formation in *Drosophila*. *Science* **339**, 443–446 (2013).
- Buchon, N., Broderick, N. A., Kuraishi, T. & Lemaître, B. *Drosophila* EGFR pathway coordinates stem cell proliferation and gut remodeling following infection. *BMC Biol.* **8**, 152 (2010).
- Hochmuth, C. E., Biteau, B., Bohmann, D. & Jasper, H. Redox regulation by Keap1 and Nrf2 controls intestinal stem cell proliferation in *Drosophila*. *Cell Stem Cell* **8**, 188–199 (2011).
- Wu, Q., Zhang, Y., Xu, J. & Shen, P. Regulation of hunger-driven behaviors by neural ribosomal S6 kinase in *Drosophila*. *Proc. Natl Acad. Sci. USA* **102**, 13289–13294 (2005).
- Biteau, B. & Jasper, H. EGF signaling regulates the proliferation of intestinal stem cells in *Drosophila*. *Development* **138**, 1045–1055 (2011).

40. Wang, L., Zeng, X., Ryoo, H. D. & Jasper, H. Integration of UPRER and oxidative stress signaling in the control of intestinal stem cell proliferation. *PLoS Genet.* **10**, e1004568 (2014).
41. Ayyaz, A., Li, H. & Jasper, H. Haemocytes control stem cell activity in the *Drosophila* intestine. *Nature Cell Biol.* **17**, 736–748 (2015).
42. Julio-Pieper, M., Flor, P. J., Dinan, T. G. & Cryan, J. F. Exciting times beyond the brain: metabotropic glutamate receptors in peripheral and non-neural tissues. *Pharmacol. Rev.* **63**, 35–58 (2011).
43. DeBerardinis, R. J. & Cheng, T. Q's next: the diverse functions of glutamine in metabolism, cell biology and cancer. *Oncogene* **29**, 313–324 (2010).
44. Guo, L., Karpac, J., Tran, S. L. & Jasper, H. PGRP-SC2 promotes gut immune homeostasis to limit commensal dysbiosis and extend lifespan. *Cell* **156**, 109–122 (2014).
45. Yan, D. Protection of the glutamate pool concentration in enteric bacteria. *Proc. Natl Acad. Sci. USA* **104**, 9475–9480 (2007).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** This work was supported by the National Institute on Aging (R01 AG028127), the National Institute on General Medical Sciences

(R01 GM100196), and by a Glenn Foundation for Medical Research postdoctoral fellowship to H.D. The Zeiss 7MP was purchased using National Institutes of Health (NIH) grant S10OD010414. We would like to thank G. Hasan, Y. Hirano, A. Toshiro, M. Parmentier, W. Lee, S. X. Hou, M. Montminy, N. Perrimon, B. Ohlstein A. DiAntonio, H. Hayashi, R. W. Daniels, K. Venkatachalam, the Vienna *Drosophila* RNAi Center, and the Bloomington Stock Center for flies, and Developmental Studies Hybridoma Bank for antibodies. We also thank Jasper laboratory members for discussion.

**Author Contributions** H.D. and H.J. designed and conceived the study. H.D. and A.A.G. performed  $\text{Ca}^{2+}$  recording and analysis, H.D. performed all other experiments in the study. H.J. and H.D. analysed data and wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.J. ([hjasper@buckinstitute.org](mailto:hjasper@buckinstitute.org)).



## METHODS

Experiments were not randomized, and investigators were not blinded to allocation during experiments and outcome assessment.

**Drosophila stocks and culture.** The following strains were obtained from the Bloomington *Drosophila* Stock Center:  $w^{1118}$ , tub::GAL80<sup>ts</sup>, UAS-mCherry, UAS-mCD8-GFP, UAS-nls-GFP, Act5C::GAL4, UAS-InR<sup>DN</sup>, UAS-InR<sup>WT</sup>, UAS-Ras<sup>V12</sup>, 20XUAS-GCaMP3 (BL32235), 20XUAS-IVS-GCaMP6s (BL42746), GAL4vGlut::GAL4<sup>OK371</sup> (BL26160), Eaat1<sup>RNAi</sup> (BL43287), Eaat1::GAL4 (BL8894), SERCA<sup>RNAi</sup> (BL44581), Eaat1<sup>OE</sup> (BL8202), SERCA<sup>Kum170</sup> (BL26700), Stim<sup>A</sup> (BL52395), Stim<sup>RNAi</sup> (BL27263), CamK1<sup>RNAi</sup> (BL26726), CamKII<sup>RNAi</sup> (BL29401), InsP3<sup>RNAi</sup> (BL25937), RyR<sup>RNAi</sup> (BL28919), PMCA<sup>RNAi</sup> (BL31572), UAS-InsP3R (BL30742), Pp2B-14D<sup>RNAi</sup> (BL40872), CanB2<sup>RNAi</sup> (BL27270 and BL38971), mGluR<sup>RNAi</sup> (BL41668), Gα<sup>RNAi</sup> (BL30735), Gβ<sup>RNAi</sup> (BL35041), UAS-CREB-b (BL9232), CREB-b<sup>RNAi</sup> (BL29332). The following lines were obtained from the Vienna *Drosophila* RNAi Center: SERCA<sup>RNAi</sup> (transformant ID 107446), Dome<sup>RNAi</sup> (transformant ID 106071), Hop<sup>RNAi</sup> (transformant ID 102830), Orai<sup>RNAi</sup> (transformant ID 12221), mGluR<sup>RNAi</sup> (transformant ID 103736), Egfr<sup>RNAi</sup> (transformant ID 43267), CanA1<sup>RNAi</sup> (transformant ID 32283). mGluR<sup>112b</sup> was a gift from M.-L. Parmentier, vGlut::GAL4<sup>CNS</sup> from A. DiAntonio, *crtc*<sup>25-3</sup> from M. Montminy, esg::GAL4 from S. Hayashi, Su(H)-Gbe::GAL4 and Gbe::GAL80 are from S. X. Hou, UAS-Hep<sup>WT</sup> from M. Mlodzik, hFlp;FRT40,tub::GAL80;tub::GAL4,UAS-GFP from B. Ohlstein. NP1::GAL4<sup>ts</sup> was from D. Ferrandon, UAS-Stim and UAS-Orai from G. Hasan, Gαq<sup>RNAi</sup> and PLCβ<sup>RNAi</sup> from D. Bohmann, UAS-xbp1<sup>spliced</sup> from P. Domingos, UAS-Notch<sup>RNAi</sup> from N. Perrimon, Su(H)-Gbe-lacZ from S. Bray, UAS-CRTC-HA and UAS-CRTC-SA-HA from Y. Hirano, CanB2<sup>KO</sup>, Pp2B-14D<sup>ACT</sup> and CanA-14F<sup>ACT</sup> from A. Toshiro and UAS-tdTomato-P2A-GCaMP5G from R.W. Daniels.

See Supplementary Information for complete list of genotypes for each figure.

Flies were cultured at 25 °C with a 12-h light–dark cycle on standard yeast/molasses-based food unless otherwise indicated (recipe for 1 l food: 750 ml H<sub>2</sub>O, 90 ml molasses, 25 g dry yeast, 6.5 g agar and 69.6 g cornmeal).

For Ca<sup>2+</sup> recording in aged flies, flies were aged at 25 °C for 30 days in cages containing populations of 20–30 flies. Flies were shifted to 29 °C 1 day before dissection and recording. Food was changed every 2 days. SERCA<sup>Kum170</sup>, FRT42D and CanB2<sup>KO</sup>, FRT42D chromosomes were recombined by neomycin selection. For MARCM clone induction, 2–3-day-old flies were heat-shocked at 37 °C for 1 h. Midguts from female flies were dissected for analysis at the time indicated.

ISC-specific knockdown was achieved by expressing double-stranded RNA (dsRNA) against mGluR using the ISC/enteroblast-specific driver esg::GAL4 combined with enteroblast-specific expression of GAL80 using Su(H)Gbe::GAL80 (ref. 40); this driver was temperature sensitive by combining it with tub::GAL80<sup>ts</sup> (Fig. 1b).

The efficiency of the mGluR<sup>RNAi</sup> construct used, VDRC 103736, was tested by qPCR (Extended Data Fig. 8), and was further enhanced by co-expressing Dicer2; no off-targets are predicted for this construct.

**Bleomycin feeding and *Ecc15* infection.** Two-to-three-day-old flies were dry starved in empty vials for 4 h before treatment. Whatman paper (Fisherbrand; Thermo Fisher Scientific) was saturated with 25 μg ml<sup>-1</sup> bleomycin (Sigma-Aldrich) dissolved in 5% sucrose. Flies were dissected for pH3 staining or for Ca<sup>2+</sup> imaging. Bacteria *Ecc15* were cultured overnight at 30 °C in LB medium. A concentrated bacterial pellet (*A*<sub>600 nm</sub> of ~200), which was centrifuged from 1-ml overnight culture media, was dissolved in 1 ml 5% sucrose in Whatman paper. Flies were infected for 6 or 24 h before dissection for pH3 staining or for Ca<sup>2+</sup> imaging.

**Calcium imaging.** Cytosolic [Ca<sup>2+</sup>] was monitored in ISCs by expressing UAS-GCaMP3, UAS-GCaMP5 or UAS-IVS-GCaMP6s under the control of esg::GAL4 combined with Su(H)Gbe::GAL80. UAS-mCherry (or tdTomato from the bi-cistronic construct with GCaMP5) was introduced as an internal control and an indicator of stem cells, and a temperature-sensitive GAL80 transgene was included to allow temporal control of transgene expression. Real-time recordings were performed 4 days after transgenes were induced at 29 °C. Similar Ca<sup>2+</sup> oscillations were observed in ISCs of age matched animals reared at 25 °C and without the temperature-sensitive GAL80 transgene (not shown). No oscillations were observed in enterocytes when GCaMP3 was expressed in these cells using NP1::GAL4 (not shown).

To visualize [Ca<sup>2+</sup>] in live ISCs, fresh intact midguts were dissected in adult-hemolymph-like (AHL) media (108 mM NaCl, 5 mM KCl, 2 mM CaCl<sub>2</sub>, 8.2 mM MgCl<sub>2</sub>, 4 mM NaHCO<sub>3</sub>, 1 mM NaH<sub>2</sub>PO<sub>4</sub>, 5 mM trehalose (Sigma), 10 mM sucrose and 5 mM HEPES, pH 7.5). Alternatively, Shields and Sang M3 insect medium (Sigma) can be used, and imaged using two-photon microscopy. The guts were slightly stretched and pinned down in a 35-mm Petri dish previously one-third filled with Sylgard (Dow Corning). Peristaltic movements were minimized by paralyzing the visceral muscle with isradipine (10 μg ml<sup>-1</sup>, Ca<sup>2+</sup> oscillations in ISCs are unaffected by isradipine; Extended Data Fig. 5f). Ca<sup>2+</sup> signals were imaged using

a Zeiss LSM 7MP two-photon microscope equipped with a Chameleon XR laser (Coherent) running at 950 nm and using a W Plan-Apochromat 20 × 1.0 numerical aperture (NA) lens. The posterior midgut area was recorded as time lapses of z-stacks comprising 17 4-μm-spaced planes every ~15 s. GCaMP3/5/6s emission was captured at 500–550 nm and mCherry/tdTomato emission at 575–610 nm.

Using automatic and manual image stabilization in Image Analyst MKII (Image Analyst Software) the movement of the gut was compensated for and mean intensities were determined in regions of interests in mean intensity projection images of the recorded z-stacks. Oscillation patterns were either calculated ‘manually’ or in an unsupervised, automated manner in Mathematica 8.0 (Wolfram Research). For the ‘manual’ calculation, oscillation frequency was determined by counting individual peaks of GCaMP3:mCherry fluorescence emission ratio observed during 12-min recordings, and average Ca<sup>2+</sup> levels were calculated for each cell as the average fluorescence ratio for the entire timeline of the experiment using Excel. Automated analysis was performed by first detecting peaks based on the second temporal derivative of the GCaMP:mCherry fluorescence emission ratio followed by a Gaussian fit of each detected oscillation spike. The local amplitude of oscillation spikes and the baseline Ca<sup>2+</sup> level between oscillations were calculated from the fit parameters. Statistical analysis using unpaired Student’s *t*-tests or ANOVA was performed in GraphPad Prism5 as indicated in the figures. Videos were generated using NIH Image J.

For pharmacological manipulations, guts were incubated in AHL media containing the respective compounds for 5–10 min and then imaged immediately. Isradipine, CdCl<sub>2</sub>, LaCl<sub>3</sub>, thapsigargin and EGTA were purchased from Sigma and 2-APB was from Tocris.

**Immunofluorescence imaging.** Intact guts were dissected, fixed and stained by immunohistochemistry following standard protocols<sup>18,40,46</sup>. Guts were dissected in PBS and fixed for 45 min at room temperature in 4% formaldehyde. All subsequent washes (1 h) and antibody incubations (4 °C overnight) were performed in PBS, 0.5% bovine serum albumin and 0.1% Triton X-100. Delta staining was performed following the methanol-heptane fixation method described previously<sup>47</sup>.

Primary antibodies and dilution: rabbit anti-β-galactosidase (Cappel), rabbit anti-dpERK (Cell 621 Signaling, 1:200), rabbit anti-peIF2α antibody (Cell Signaling: 3597, 1:150), rabbit anti-622 VGlut (gift from A. DiAntonio), rat anti-Delta (gift from M. D. Rand, 1:1,000), anti-phospho-histone H3 Ser10 (Upstate, 1:1,000), and anti-prospero, anti-armadillo, anti-β-galactosidase and anti-Delta (Developmental Studies Hybridoma Bank), 1:300, 1:250, 1:100, and 1:100, respectively. Alexa-Fluor-647-anti-HRP and all fluorescent secondary antibodies were from Jackson ImmunoResearch. Phalloidin (Invitrogen, 1:400) was used to label the visceral muscle in the midgut in Fig. 2c. DAPI was used to stain DNA. All images were taken on a Zeiss LSM 710 confocal microscope and processed using Adobe Photoshop, Illustrator and Image J.

For gut cross sections, guts were fixed using methanol/heptane and briefly washed twice in PBS. The posterior midgut was cut into sections of 0.2–1 mm in thickness by a fine steel blade. The gut pieces were flipped and mounted sagittally.

**In situ hybridization.** The protocol was adapted from ref. 48 using digoxigenin (DIG)-labelled RNA probes detected by NBT/BCIP assay (DIG nucleic acid detection kit, Roche). The following primers were used to generate RNA probes for *Eaat1*: forward 5′-AGATTGCTACCAGAGCCATTAC-3′ and reverse 5′-GCACCAACAGGAATGACAAATC-3′.

**Gut growth measurement.** The width of the ‘distal hairpin’ was measured as an indicator of gut growth. The distal hairpin region is adjacent to the copper cells in the posterior midgut. The width of this region was measured using ZEN software and reported as arbitrary pixel values. Gut length (pixel) was measured by ImageJ. For cell density measurement, a 70 × 70 μm area in the posterior midgut was chosen and cell number was counted based on DAPI staining. At least 10 guts of each genotype were counted.

**CAFÉ assay and food intake experiment.** Food naive control flies were water starved for 2 days and refed with the CAFÉ assay based on ref. 49 with minor modification. In brief, capillaries filled with liquid food (5% sucrose or 5% sucrose plus 1% glutamate) were marked with lines before and after 4 h feeding. The length in between (in cm) was calculated and compared in control food and glutamate supplemented food. At least three sets of biological replicates were compared for each food condition.

For coloured food ingestion experiments, groups of ddH<sub>2</sub>O starved food naive control flies were transferred onto fresh food medium as indicated containing 2.5% (w/v) blue food dye (FD&C Blue Dye no. 1). Images were taken by Zeiss dissection microscope after 4 h ingestion.

**Glutamate feeding.** The feeding regimen is based on ref. 1 with minor modification. In brief, black pupae raised on regular cornmeal/molasses food were collected and transferred to new empty vials with ddH<sub>2</sub>O soaked Whatman paper. To allow for transgene induction, flies were shifted to 29 °C 2 days before hatching. After eclosion, food-naive flies were water starved (soaked Whatman paper) for another

2 days at 29°C. Flies were then refed protein restricted food supplemented with various amounts of L-Glu for the periods indicated in the figures. Restricted food recipe: 5% sucrose, 1% agar and 0.1% yeast. L-Glu (G5889), L-Gly (G2879), L-Asp (A9256), L-Ser (S4500), L-Cys (168149) and L-Gln (90114C) were purchased from Sigma. Amino acids were dissolved in water and mixed with the boiled restricted food at different concentrations. Food was changed daily. For pH3 staining and gut growth, flies were dissected, fixed and stained by standard protocols. For Ca<sup>2+</sup> imaging after L-Glu treatment, food-naive flies were ddH<sub>2</sub>O starved for 2 days at 29°C and recorded in regular AHL media or AHL plus 1% glutamate.

**Glutamate microinjection.** The injection protocol was as described previously<sup>41</sup>. In brief, thin sterile glass capillaries were filled with the indicated solutions and 50 nl were injected into the thorax of female adults using a NanojectII apparatus (Drummond). Fluid dispersed throughout the animal as indicated by blue food dye (FD&C Blue Dye no.1) mixed into the solutions. To study whether circulating glutamate can influence ISC proliferation in the starvation/refeeding paradigm, food-naive flies were starved on water for 2 days and injected with PBS or 1% glutamate in PBS. The injected flies were recovered and raised in restricted food condition (0.1% yeast, 5% sucrose and 1% agar) for the indicated time periods before dissection.

**Statistical methods.** Statistical analysis was performed using GraphPad Prism5 and Microsoft Excel. Statistical methods used and sample sizes are listed in the figure legends. One-way ANOVA with post-hoc Dunnett test performed for Figs 3a, b, 4b and Extended Data Figs 2a (left), 2e, 5c, e, 6d, f, 7c, 8b, 8f, h, i, 9a, c, d, f, i, k and 10b, i, j, l. One-way ANOVA with post-hoc Tukey test was used for Figs 3f, 4b (right), Extended Data Figs 2a (right), 3d (middle), 3f, 9e and 10h. *P* values for other figures were determined by Student's *t*-test. Sample sizes were chosen empirically based on observed effect sizes.

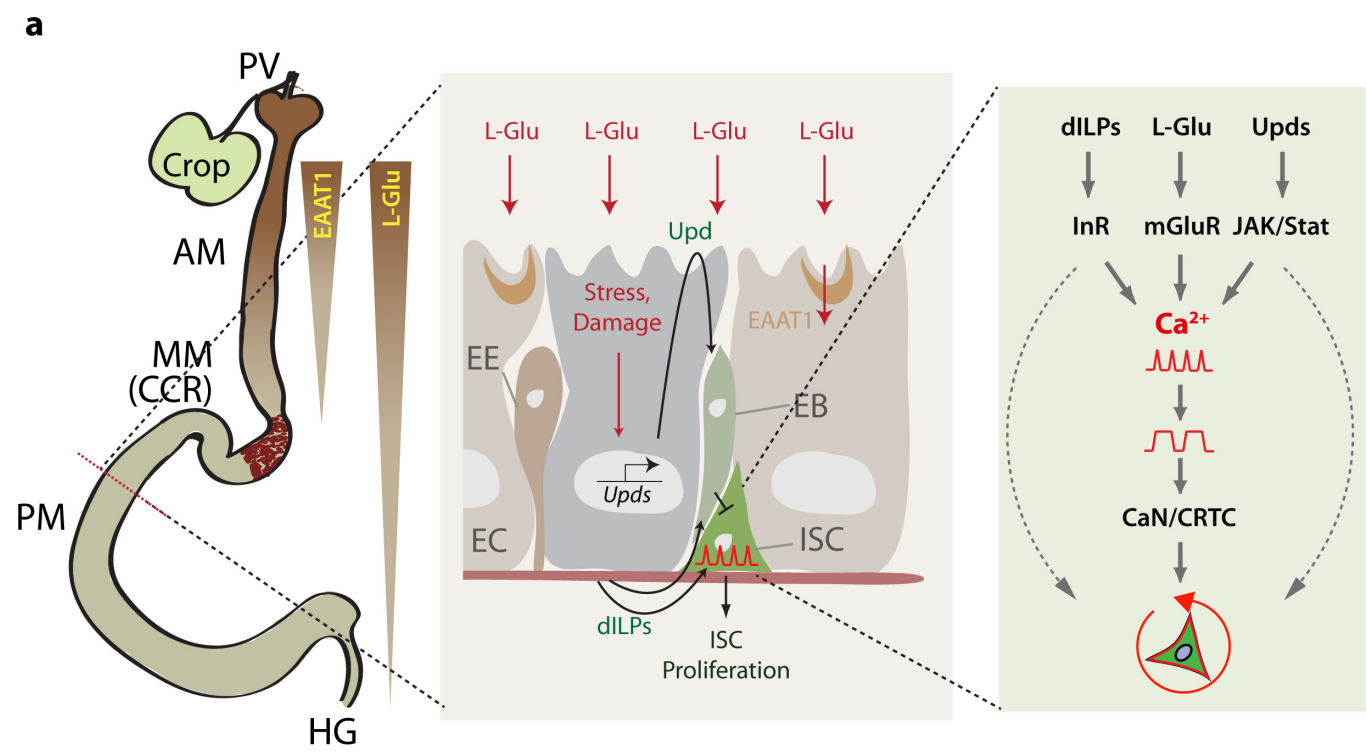
**qRT-PCR analysis of gene expression.** Total RNA was extracted from dissected guts (10 per sample) or heads (4 per sample) using Trizol, and cDNA was synthesized using Superscript III (Invitrogen). qPCR was performed in triplicate for each sample from two independent (Extended Data Fig. 9m) or three independent (Extended Data Fig. 3c, f) experiments using SYBR Green on a Biorad IQ5 system. Expression was normalized with *Act5C*.

**Primer sequences for RT-PCR.** The following primers pairs were used: *Act5C* (forward): 5'-CTCGCCACTTGCGTTTACAGT-3', *Act5C* (reverse):

5'-TCCATATCGTCCCAGTTGGTC-3'; *CaMKI* (forward): 5'-CGCTACAGTGATCCGACAAA-3', *CaMKI* (reverse): 5'-CTGTTGGAGTGGT AGAGTCTTG-3'; *CaMKII* (forward): 5'-GTGGATTGTCAACGCGAAC-3', *CaMKII* (reverse): 5'-GTCGTAAGTATGGCTCCCTTTAG-3'; *Stim* (forward): 5'-CCAGAGAGGCTTGTGAGAAAT-3', *Stim* (reverse): 5'-CGACCGATCCACATCATCAATA-3'; *CanB2* (forward): 5'-GTTCGATGAGTTCT GC TCTGT-3', *CanB2* (reverse): 5'-GAAGGATGCCGACGTGTATAA-3'; *InsP3R* (forward): 5'-TACCTCGGGAGTCTGGAATTAG-3', *InsP3R* (reverse): 5'-CTCCCAGTAAGCAGAGAGTAGA-3'; *Eaat1* (forward): 5'-CCAGACGGAGA GCATTGATAAG-3', *Eaat1* (reverse): 5'-GGAACATGGTCGACTGATGA-3'; *mGluR* (forward): 5'-CGTCAGTGGCTCTTGTATGT-3', *mGluR* (reverse): 5'-GTAGACGGTGCTGTTTCATAGTC-3'; *Upd3* (forward): 5'-CCCAGCC AACGATTTTATG-3', *Upd3* (reverse): 5'-TGTACCGCTCCGGCTAC-3'.

46. Biteau, B., Hochmuth, C. E. & Jasper, H. JNK activity in somatic stem cells causes loss of tissue homeostasis in the aging *Drosophila* gut. *Cell Stem Cell* **3**, 442–455 (2008).
47. Lin, G., Xu, N. & Xi, R. Paracrine Wingless signalling controls self-renewal of *Drosophila* intestinal stem cells. *Nature* **455**, 1119–1123 (2008).
48. Takashima, S. et al. Development of the *Drosophila* entero-endocrine lineage and its specification by the Notch signaling pathway. *Dev. Biol.* **353**, 161–172 (2011).
49. Ja, W. W. et al. Prandiology of *Drosophila* and the CAFE assay. *Proc. Natl Acad. Sci. USA* **104**, 8253–8256 (2007).
50. Grandison, R. C., Piper, M. D. & Partridge, L. Amino-acid imbalance explains extension of lifespan by dietary restriction in *Drosophila*. *Nature* **462**, 1061–1064 (2009).
51. Lee, W. C. & Micchelli, C. A. Development and characterization of a chemically defined food for *Drosophila*. *PLoS ONE* **8**, e67308 (2013).
52. Olds, W. H. & Xu, T. Regulation of food intake by mechanosensory ion channels in enteric neurons. *Elife* **3**, 04402 (2014).
53. Cognigni, P., Bailey, A. P. & Miguel-Aliaga, I. Enteric neurons and systemic signals couple nutritional and reproductive status with intestinal homeostasis. *Cell Metab.* **13**, 92–104 (2011).
54. Sanyal, S., Jennings, T., Dowse, H. & Ramaswami, M. Conditional mutations in SERCA, the Sarco-endoplasmic reticulum Ca<sup>2+</sup>-ATPase, alter heart rate and rhythmicity in *Drosophila*. *J. Comp. Physiol. B* **176**, 253–263 (2006).
55. Ha, E. M. et al. Regulation of DUOX by the Gαq-phospholipase Cβ-Ca<sup>2+</sup> pathway in *Drosophila* gut immunity. *Dev. Cell* **16**, 386–397 (2009).

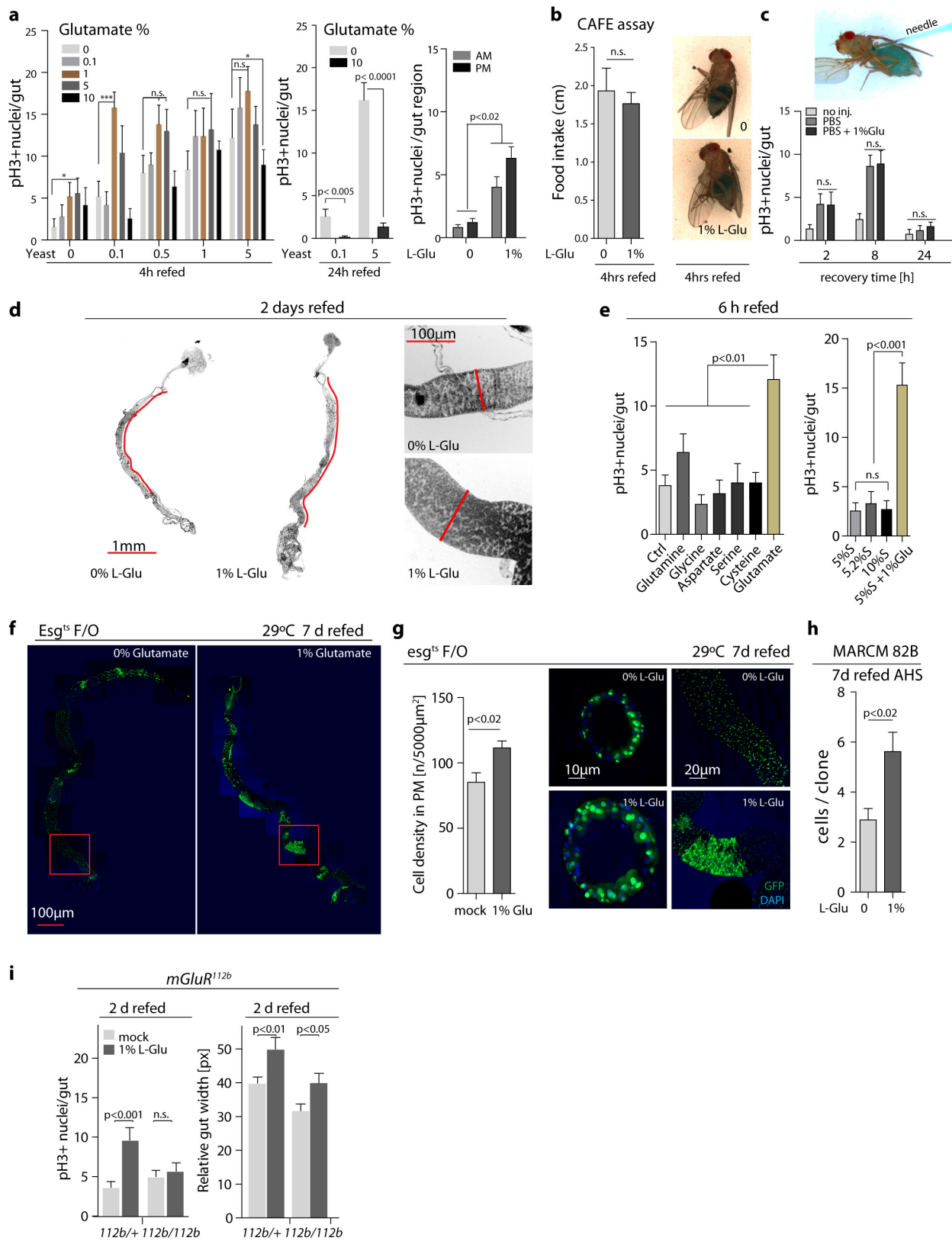




**b**

Category	<i>Drosophila</i> gene name	Category	<i>Drosophila</i> gene name
GPCR pathway components	<i>GaQ</i> ( <i>Galpha49B</i> )	$\text{Ca}^{2+}$ effectors	<i>Pp2B-14D</i>
	<i>GaF</i> ( <i>Galpha73B</i> )		<i>CanA-14F</i>
SOCE channels	<i>Gal</i> ( <i>G-ialpha65A</i> )		<i>CanA1</i>
	<i>GaO</i> ( <i>G-oalpha47A</i> )		<i>CanB</i>
	<i>GaS</i> ( <i>G-salpha60A</i> )		<i>CanB2</i>
	<i>Gβ</i> ( <i>Gβ13F</i> )		<i>CRTC (TORC)</i>
	<i>Gγ1</i> ( <i>Ggamma1</i> )		<i>CaMKI</i>
	<i>PLCbβ</i> ( <i>NorpA</i> )		<i>CaMKII</i>
	<i>SERCA</i> ( <i>Ca-P60A</i> )	VGCCs (Voltage gated $\text{Ca}^{2+}$ channels)	<i>Cac</i> ( <i>cacophony</i> )
	<i>Orai</i> ( <i>olf186-F</i> )		<i>Ca-alpha1T</i>
	<i>Stim</i>		<i>Ca-alpha1D</i>
	<i>InsP<sub>3</sub>R</i> ( <i>Itp-r83A</i> )		<i>Ca-beta</i>
	<i>RyR</i> ( <i>Rya-r44F</i> )		

**Extended Data Figure 1 | Integration of stress and dietary signals by  $\text{Ca}^{2+}$  signalling.** **a**, We propose that  $\text{Ca}^{2+}$  signalling integrates and transduces nutritional and stress signals from the environment and from systemic and local paracrine sources to respond appropriately to tissue needs. In homeostatic conditions, ISCs are largely quiescent. Glutamate derived from protein in the food is absorbed by enterocytes through Eaat1, which is expressed throughout the gut, but at relatively higher levels in the anterior midgut. We propose that excessive dietary L-Glu is not efficiently resorbed by enterocytes and activates mGluR in ISCs. The L-Glu signal is integrated with local and systemic stress and growth factor signals to promote ISC proliferation in a CaN/CRTC dependent manner. CaN/CRTC signalling is required and sufficient downstream of InR and Jak/Stat signalling to induce ISC proliferation. We propose that Crtc acts together with other transcription factors regulated by InR and Jak/Stat signalling. **b**,  $\text{Ca}^{2+}$  signalling components in *Drosophila* (Extended Data Fig.1b).

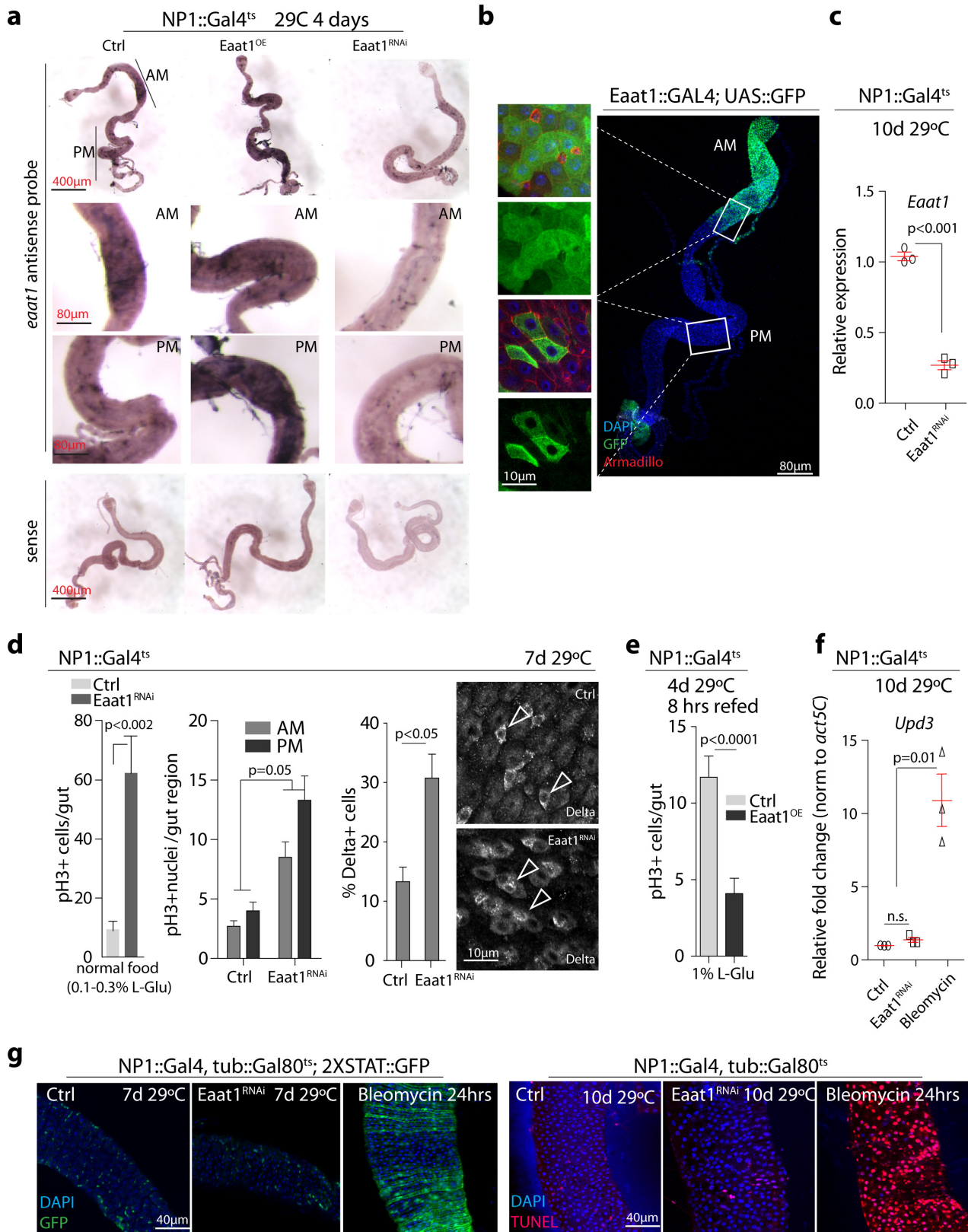


Extended Data Figure 2 | See next page for caption.



**Extended Data Figure 2 | Dietary L-Glu stimulates mGluR-dependent gut growth.** **a**, Left, mitotic figures in intestines refeed with a range of L-Glu (w/vol%) in food with varying yeast concentrations for 4 h. Flies were starved after eclosing for 48 h, and then refeed with food containing varying concentrations of yeast as only amino acid source, but supplemented with varying concentrations of L-Glu (between 0.1 and 10% L-Glu in food containing between 0.1% and 5% yeast; note that the L-Glu concentration in standard fly food is about 0.1–0.3%; refs 50, 51). Middle, 10% L-Glu refeeding inhibits ISC proliferation. Mitotic figures were quantified 24 h after flies were refeed with yeast-enriched (5%) or yeast-restricted (0.1%) food supplemented with 10% glutamate. Blue food dye was included in the food to monitor food intake. Right, distribution of mitotic ISCs along the gastrointestinal tract after L-Glu refeeding. Number of proliferating ISCs (pH3<sup>+</sup>) in anterior and posterior midguts (AM and PM, respectively) is quantified. **b**, Food intake is not affected by changing the L-Glu concentration in the food. CAFÉ assay<sup>49</sup> and ingestion of coloured food are shown after 4 h of refeeding. **c**, Injection of L-Glu fails to promote ISC proliferation. Food-naïve flies were starved for 2 days before injection with the indicated solutions. Injected or non-injected flies were examined after recovery for the indicated time points. A typical example of injected fly was shown on the top. Blue food dye was mixed into the injected solutions to monitor its distribution throughout the body. **d**, Representative images of guts after refeeding. The red lines indicate the length and width measurements used to quantify relative gut size. **e**, Left, ISC proliferation in animals refeed sucrose solution supplemented with indicated amino acids (1% w/vol final). The number of pH3<sup>+</sup> ISCs was determined 6 h after refeeding. Right, ISC proliferation in animals refeed isocaloric sucrose-only solution (5.2% sucrose, the calorie content of 0.2%

sucrose is equivalent to 1% glutamate) or refeed excess sucrose (10%). **f**, Feeding L-Glu for 7 days promotes growth of *esg*<sup>ts</sup> FLP-out (F/O) clones. Note large clones in L-Glu-fed intestines (box). Genotype of *esg*<sup>ts</sup> F/O: *esg*::GAL4, UAS-FLP, tub::GAL80<sup>ts</sup>; UAS-nls-GFP; tub>FRT-CD2-FRT>GAL4. **g**, Increased cell density, lineage growth and intestinal diameter in L-Glu-treated guts. Cell density in the posterior midgut was determined in *esg*<sup>ts</sup> F/O flies exposed to mock or L-Glu-supplemented food for 7 days (*esg*<sup>ts</sup> F/O allows lineage tracing from ISCs). Genotype: *esg*::GAL4, UAS-FLP, UAS-GFP, Act>FRT-stop-FRT>GAL4; tub::GAL80<sup>ts</sup>. Cross sections through the posterior midgut and overview images are shown highlighting the increase in intestinal diameter and in clonal growth after L-Glu supplementation. Clones were analysed 7 days after clone induction at 29 °C. **h**, L-Glu supplementation promotes growth of ISC lineages. ISC lineages were marked by MARCM (genotype: *hs*-FLP, UAS-GFP; tub::GAL4, FRT82B, tub::GAL80/FRT82B). Cell numbers in GFP<sup>+</sup> ISC clones were counted 7 days after heat-shock induction and refeeding. **i**, Related to Fig. 1b. Mitotic figures and gut width measured after L-Glu refeeding. *mGluR*<sup>112b</sup> is a null allele of *mGluR*. For pH3 number, *n* = 12 per condition, for gut width measurement, *n* = 8 per condition. Data and s.e.m. are shown. *P* values from Student's *t*-test (**a** (middle), **b**, **c**, **g**, **h** and **i**) and one-way ANOVA (**a** (left and right) and **e**). The sample size (per condition) is as follows: *n* = 8 for **a** (left), *n* = 7 for **a** (middle), *n* = 11 for **a** (right), *n* = 12 for **b** and **c**, *n* = 9 for **d**, *n* = 11 for **e** (left), *n* = 6 for **e** (right), *n* = 6 for **g**, *n* = 14 for **i** (left) and *n* = 9 for **i** (right). For **h**, clones (*n* = 32 for control and *n* = 42 for glutamate food) from 5 guts were analysed. Data are representative of 3 (**a**, **e** and **g**) and 2 (**b**, **c**, **h** and **i**) independently performed experiments.

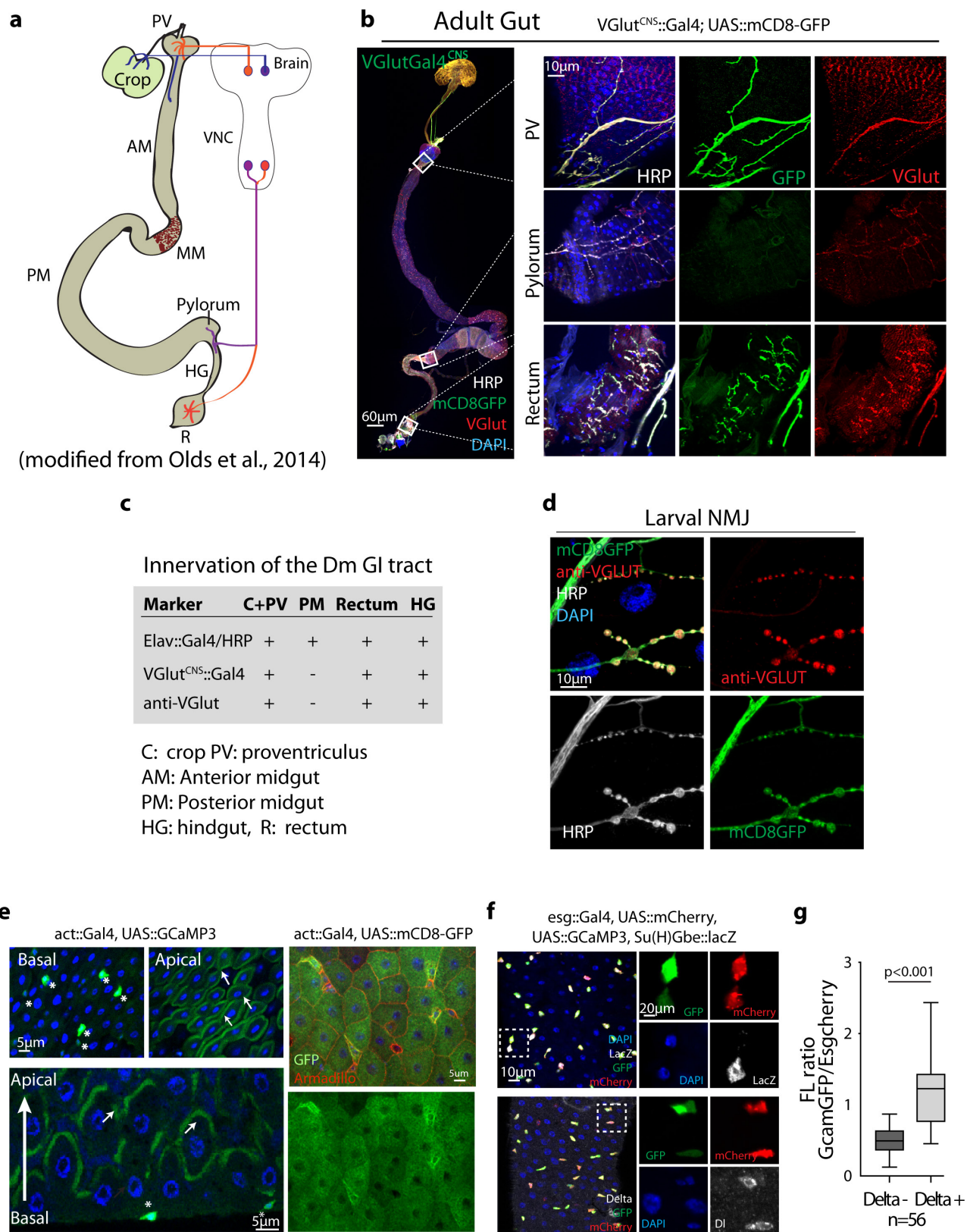


Extended Data Figure 3 | See next page for caption.



**Extended Data Figure 3 | ISC regulation by *Eaat1*.** **a**, *Eaat1* transcript is enriched in the anterior midgut. *In situ* hybridization with antisense RNA probes against *Eaat1* (detected using NBT/BCIP; sense probes shown in bottom panels). Guts overexpressing *Eaat1* in enterocytes (*Eaat1*<sup>OE</sup>), or depleted of *Eaat1* in enterocytes (*Eaat1*<sup>RNAi</sup>) used as positive/negative controls. **b**, Expression pattern of *Eaat1*::GAL4 in midgut. Expression of *Eaat1*::GAL4 > UAS-GFP (green) in anterior and posterior midguts shown in higher magnification in the bottom panels. Cell membranes highlighted by anti-armadillo staining (red). **c**, Knockdown efficiency of *Eaat1*<sup>RNAi</sup> determined by qRT-PCR. *Eaat1*<sup>RNAi</sup> (BL43287) was used to knockdown *Eaat1* in the gut using NP1::GAL4, tub::GAL80<sup>ts</sup> (29 °C for 10 days). NP1::GAL4<sup>ts</sup> (NP1::GAL4; tub::GAL80<sup>ts</sup>) drives expression in enterocytes throughout the gut when flies are shifted to 29 °C. Ten guts were pooled for RNA extraction and three independent groups were repeated for evaluation. **d**, Knocking down *Eaat1* in enterocytes promotes ISC proliferation. Distribution of mitotic ISCs and percentage of D1<sup>+</sup> cells in intestines of flies in which *Eaat1* was knocked down in enterocytes shown in the middle panels. Representative image from confocal microscopy is shown on the right. Arrowheads point to select ISCs identified by anti-D1 staining (white). **e**, Overexpressing *Eaat1* in enterocyte suppresses L-Glu-mediated ISC proliferation. *Eaat1* was overexpressed in enterocytes using NP1::GAL4<sup>ts</sup> (NP1::GAL4; tub::GAL80<sup>ts</sup>) at 29 °C. Flies were shifted to 29 °C before hatching, then

maintained on normal food (0.1–0.3% L-Glu) for 4 days. Flies were then starved (with water) for another 2 days. Mitotic index was determined in intestines of flies refed with 1% glutamate for 6 h as shown in **a**. **f**, *upd3* transcript in whole guts quantified by qRT-PCR (transcript levels normalized to *Act5C*). **g**, Knocking down *Eaat1* in enterocytes by NP1::GAL4<sup>ts</sup> does not increase Jak/Stat signalling activity in the posterior midgut. 2×STAT-GFP is an activity reporter for Jak/Stat signalling (green). Apoptosis in enterocytes is indicated by TUNEL staining (bottom, apoptotic nuclei in red). As a positive control, bleomycin (25 µg ml<sup>-1</sup>) treatment strongly induces 2×STAT-GFP expression in visceral muscle and epithelium, results in widespread TUNEL-positive nuclei in the intestinal epithelium, and substantially increases *upd3* transcripts in the posterior midgut. Data are mean and s.e.m. *P* values are from ANOVA (**d** (middle) and **f**) and Student's *t*-test (**c**, **d** (left and right) and **e**). The sample size is *n* = 3 independent samples for each condition in **c**, *n* = 12 guts for each condition (mitotic figures), and *n* = 14 guts for control and *n* = 24 guts for *Eaat1*<sup>RNAi</sup> (percentage of D1<sup>+</sup> cells) in **d**; *n* = 15 for control and *n* = 21 guts for *Eaat1*<sup>OE</sup> in **e**; *n* = 3 independent samples for each condition in **f**. For **d** and **e**, a representative experiment is shown (three biological replicates). For **a**, **b** and **g**, representative images from a representative of two independent experiments are shown (**a**: *n* = 8 for each condition; **b**: *n* = 12; **g** (left): *n* = 12 per condition, and **g** (right): *n* = 9 per condition).

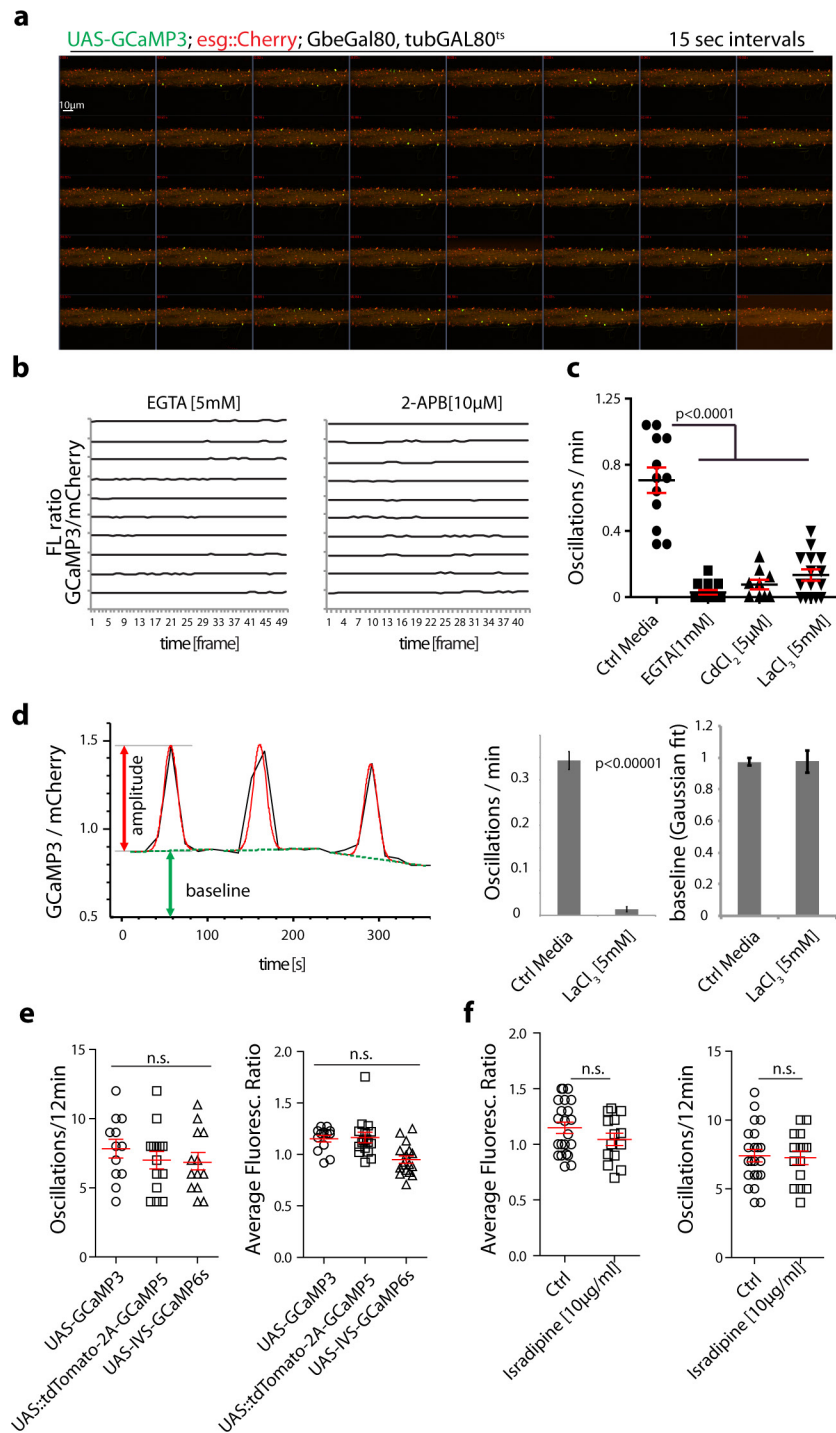


Extended Data Figure 4 | See next page for caption.

**Extended Data Figure 4 | Neuronal projections and GCaMP3 fluorescence pattern in the intestinal epithelium.** **a**, Schematic showing neuronal projections in adult *Drosophila* gastrointestinal tract (modified from ref. 52). Only three segments are innervated by neurons from the brain: crop, proventriculus (PV) and foregut, hindgut (HG) pylorus, and the rectum (R). **b**, No glutamatergic neurites are observed in the posterior midgut<sup>52,53</sup>. All neurites are stained by anti-horseradish-peroxidase (HRP) staining in white. mCD8-GFP driven by GAL4vGlut::GAL4<sup>CNS</sup> (a glutamatergic neuron driver readily expressed in adults) labels all glutamatergic neurons. Glutamatergic neurons are also detected by VGlut (vesicular glutamate transporter) immunostaining. Higher magnification is shown on the right. **c**, Summary of neurites innervating the *Drosophila* gastrointestinal tract (see also ref. 53). While L-Glu in the diet may stimulate glutamatergic neurons innervating the intestine, and a role for these neurons in stimulating ISC activity under certain circumstances cannot be ruled out, the widespread induction of ISC proliferation throughout the gut, as well as the ISC-specific requirement for mGluR suggests that the locally restricted glutamatergic innervation is not directly involved in stimulating ISC proliferation. **d**, GAL4vGlut::GAL4<sup>CNS</sup>>mCD8-GFP and anti-vGlut stain the presynaptic neurons of third instar larval neuromuscular junctions. HRP (white) stains all neurites. **e**, Detection of Ca<sup>2+</sup> levels in the

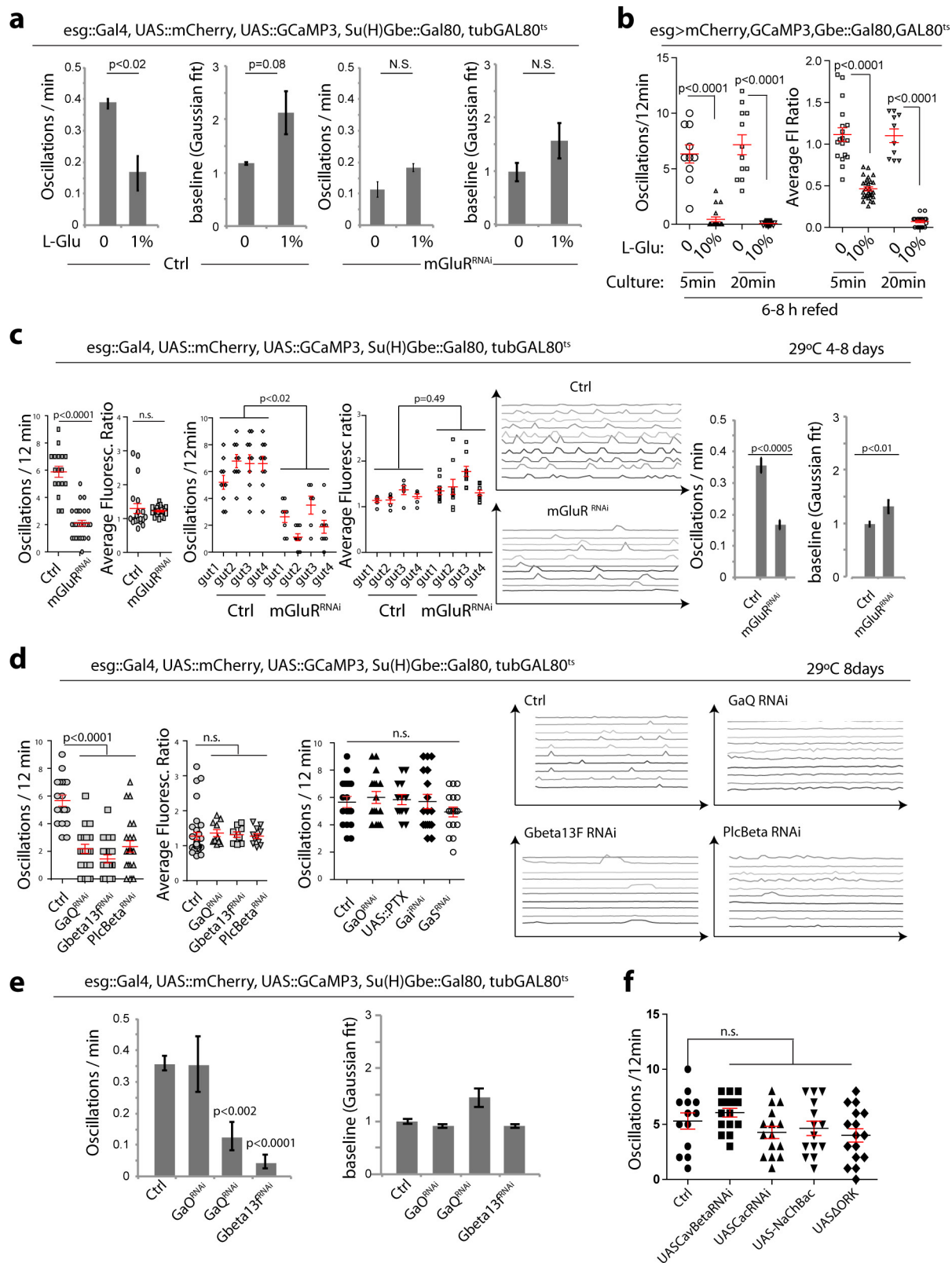
*Drosophila* posterior midgut (fixed tissue). Ubiquitous expression of UAS-GCaMP3 (green) in the posterior midgut using Actin5C::GAL4 reveals high Ca<sup>2+</sup> levels in subsets of small, basally located cells (likely to be ISCs and enteroblasts; asterisks), and at the brushed border of enterocytes (arrows). Superficial view in top panels, sagittal view in bottom panels. Right, Actin5C::GAL4 drives expression of mCD8-GFP homogeneously in the posterior midgut. Armadillo (red) marks cell boundaries. In enterocytes, Ca<sup>2+</sup> is enriched in the microvilli (marked with arrows), potentially to facilitate Ca<sup>2+</sup> dependent absorption and innate immune processes<sup>4</sup>. **f**, Ca<sup>2+</sup> concentration is higher in ISCs than in enteroblasts. ISCs and enteroblasts are labelled by expression of mCherry (red) under the control of *esg*::GAL4. Ca<sup>2+</sup> in these cells is detected using GCaMP3 (green). Su(H)Gbe-lacZ marks enteroblasts (white in top panels) and D1 marks ISCs (white in bottom panels). **g**, Quantification of relative Ca<sup>2+</sup> levels in ISCs versus enteroblasts measured as fluorescence ratio between GCaMP3 and mCherry. Note the high variability of [Ca<sup>2+</sup>] in ISCs. Data are mean and s.e.m. *P* values from Student's *t*-test. *n* = 56 cell pairs from 5 different guts. Four independent experiments were performed. One representative image from 5 flies in a single experiment (two independent duplicates) is shown in **b**, **e** and **f**. One representative image from 6 larvae in a single experiment (two independent duplicates) is shown in **d**.





**Extended Data Figure 5 |  $\text{Ca}^{2+}$  oscillations in ISCs.** **a**, Typical *ex vivo* recording of  $\text{Ca}^{2+}$  oscillations in ISCs from young flies. GfCaMP3 (green) and mCherry (red) are expressed specifically in ISCs. A total of 56 frames with 15-s intervals are shown (see Methods for details). Frames were exported from ZEN software. Genotype: UAS-GfCaMP3; esg::GAL4, UAS-mCherry; Su(H)Gbe::GAL80, tub::GAL80<sup>ts</sup>. **b**, Traces of  $\text{Ca}^{2+}$  oscillations in ISCs of intestines incubated with EGTA (5 mM) or 2-APB (10  $\mu\text{M}$ ) for 5–10 min and then recorded immediately. **c**,  $\text{Ca}^{2+}$  oscillation frequency of ISCs incubated in AHL with or without EGTA,  $\text{CdCl}_2$  (5  $\mu\text{M}$ ) or  $\text{LaCl}_3$  (5 mM). Values for individual ISCs (collected from 3–4 different guts) are plotted. **d**,  $\text{Ca}^{2+}$  oscillation in ISCs is inhibited by  $\text{LaCl}_3$ . Scheme of oscillation parameters is shown on the left. The local amplitude of oscillation spikes and the baseline  $\text{Ca}^{2+}$  level between oscillations were derived from Gaussian fits on detected oscillation spikes. Average

oscillation frequencies and average baselines are shown. **e**, Observed  $\text{Ca}^{2+}$  oscillation pattern in ISCs is independent of the genetic reporter system used. Oscillation frequency and average fluorescence ratio are compared between different genetic Calcium reporters: UAS-GfCaMP3, UAS-tdTomato-P2A-GfCaMP5 and UAS-IVS-GfCaMP6s. Genetic reporters are expressed specifically in ISCs using esg::GAL4<sup>ts</sup>; Su(H)Gbe::GAL80. **f**,  $\text{Ca}^{2+}$  oscillation pattern in ISCs of guts incubated with or without isradipine (10  $\mu\text{g ml}^{-1}$  incubation in AHL medium). Isradipine inhibits L-type VGCCs to paralyze visceral muscles. Data are mean and s.e.m. *P* value from ANOVA (**c** and **e**) or Student's *t*-test (**d**, **f**). Individual ISCs pooled from 3–4 guts were plotted in **c**, **e** and **f**. The sample size for **c** is  $n = 12, 8, 9$  and  $14$ ; for **e**,  $n = 12, 16, 13, 14, 20$  and  $24$ ; and for **f**,  $n = 22, 14, 25$  and  $22$  (from left to right for each panel). Two (**b–d**) and three (**e**, **f**) independent experiments were performed.

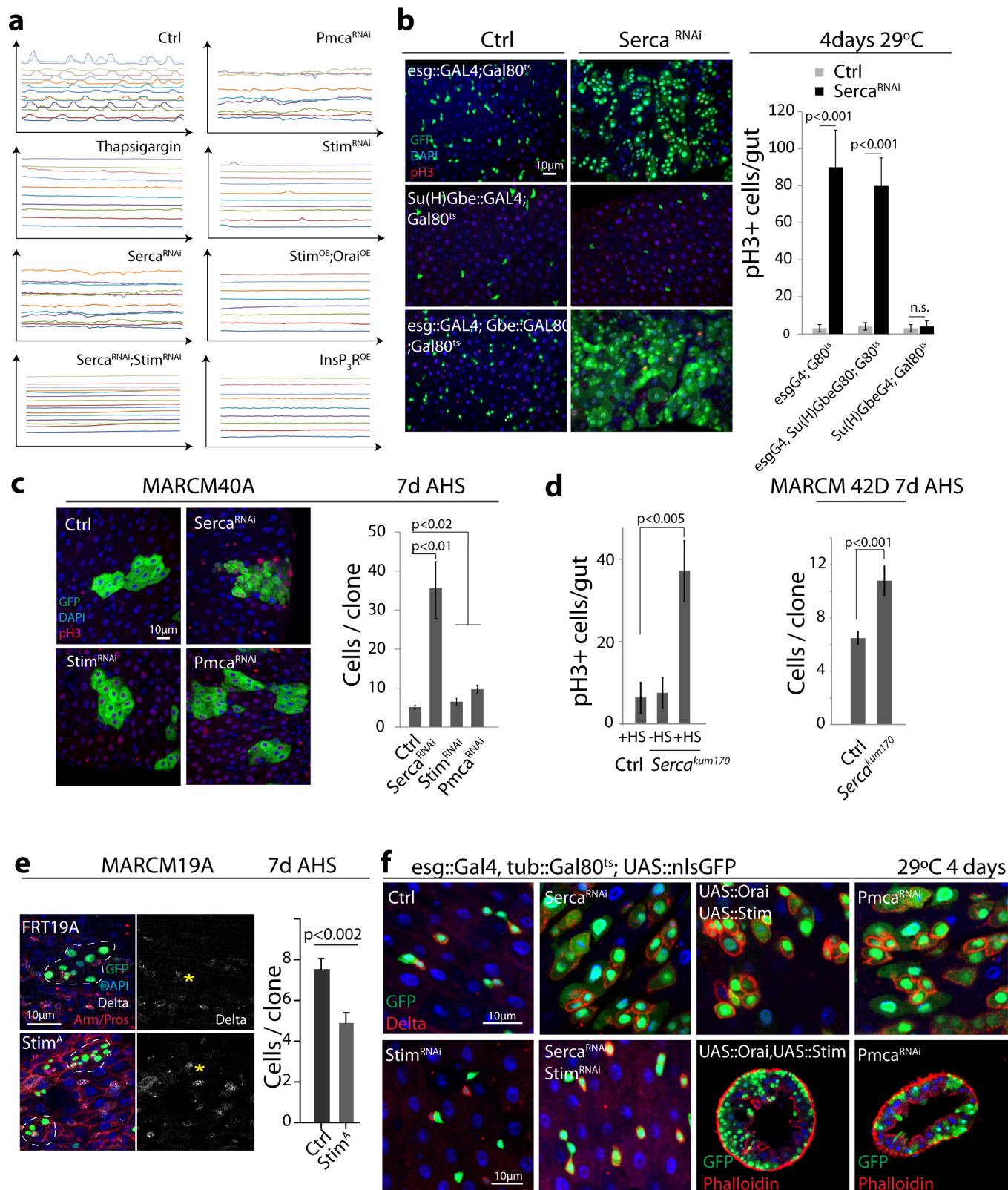


Extended Data Figure 6 | See next page for caption.

**Extended Data Figure 6 | Glutamate feeding modulates  $\text{Ca}^{2+}$  oscillation pattern through the mGluR/ $\text{G}\alpha\text{q}$ /PLC $\beta$  pathway.** **a**, Automated quantification of oscillation parameters in recordings from wild-type and mGluR<sup>RNAi</sup> expressing intestines refed L-Glu containing or control food (compare with Fig. 3a). Average oscillation frequencies and average baselines are shown. **b**, 10% L-Glu feeding inhibits  $\text{Ca}^{2+}$  oscillations.  $\text{Ca}^{2+}$  oscillations recorded in animals after 6–8 h refeeding with 10% L-Glu and incubation of the intestine in AHL supplemented with 10% L-Glu. Prolonged incubation (20 min) in 10% L-Glu causes further decrease of average GCaMP3 fluorescence. **c**, Manual and automated quantification of oscillation parameters on recordings from individual guts of the indicated genotypes. Values for individual cells from 3–4 guts are plotted on the left. Values for individual cells from single guts are shown in the third and fourth panels. Typical traces of recordings for individual cells are shown in the middle and the corresponding oscillation parameters calculated automatically by Gaussian fits in Mathematica 8.0 are shown on the right. Average oscillation frequencies and average baselines are shown. **d**,  $\text{G}\alpha\text{q}$ , but not other  $\text{G}\alpha$  subunits ( $\text{G}\alpha\text{O}$ ,  $\text{G}\alpha\text{S}$  or  $\text{G}\alpha\text{i}$ ) is required for  $\text{Ca}^{2+}$  oscillations in ISCs. While knockdown of  $\text{G}\alpha\text{q}$ ,  $\text{G}\beta 13\text{F}$  and PLC $\beta$

impairs  $\text{Ca}^{2+}$  oscillations, oscillation frequency was not affected when  $\text{G}\alpha\text{O}$ ,  $\text{G}\alpha\text{S}$  or  $\text{G}\alpha\text{i}$  were knocked down in ISCs for 8 days at 29 °C (driver genotype: UAS-GCaMP3; esg::GAL4, UAS-mCherry, Su(H)Gbe::GAL80, tub::GAL80<sup>ts</sup>). Pertussis toxin (UAS-PTX) was used to inhibit  $\text{G}\alpha\text{q}$  specifically. Typical traces are shown on the right. **e**, Average oscillation frequencies and average baselines from Gaussian fits are shown. **f**, Voltage-gated calcium channel (VGCC) components Cav $\beta$  or Cac are not required for  $\text{Ca}^{2+}$  oscillations in ISCs. Oscillations are also not affected in ISCs expressing molecules expected to affect plasma membrane potential ( $\Delta\text{Ork1}$  (hyperpolarizing) and UAS-NaChBac (hypo-polarizing)). Driver genotype: UAS-GCaMP3; esg::GAL4, UAS-mCherry; Su(H)Gbe::GAL80, tub::GAL80<sup>ts</sup>. Data are mean and s.e.m. *P* values are from ANOVA (**c** (right), **d** and **f**) and Student's *t*-test (**a**, **b**, **c** (left) and **e**). Individual ISCs pooled from 3–4 guts are plotted in **b–d** and **f**. The sample size for **a** and **e** is *n* = 3; for **b** *n* = 10, 8, 12, 17, 15, 20, 10 and 11; for **c**, *n* = 16, 20, 17 and 12; for **d**, *n* = 15, 18, 13, 22, 22, 10, 8 and 9; and for **f**, *n* = 13, 15, 15, 14 and 16 (from left to right for all panels). For **c**, values for individual cells from single guts are shown in the third and fourth panels. Data are representative of 2 (**d–f**) and 3 (**a–c**) independent experiments.

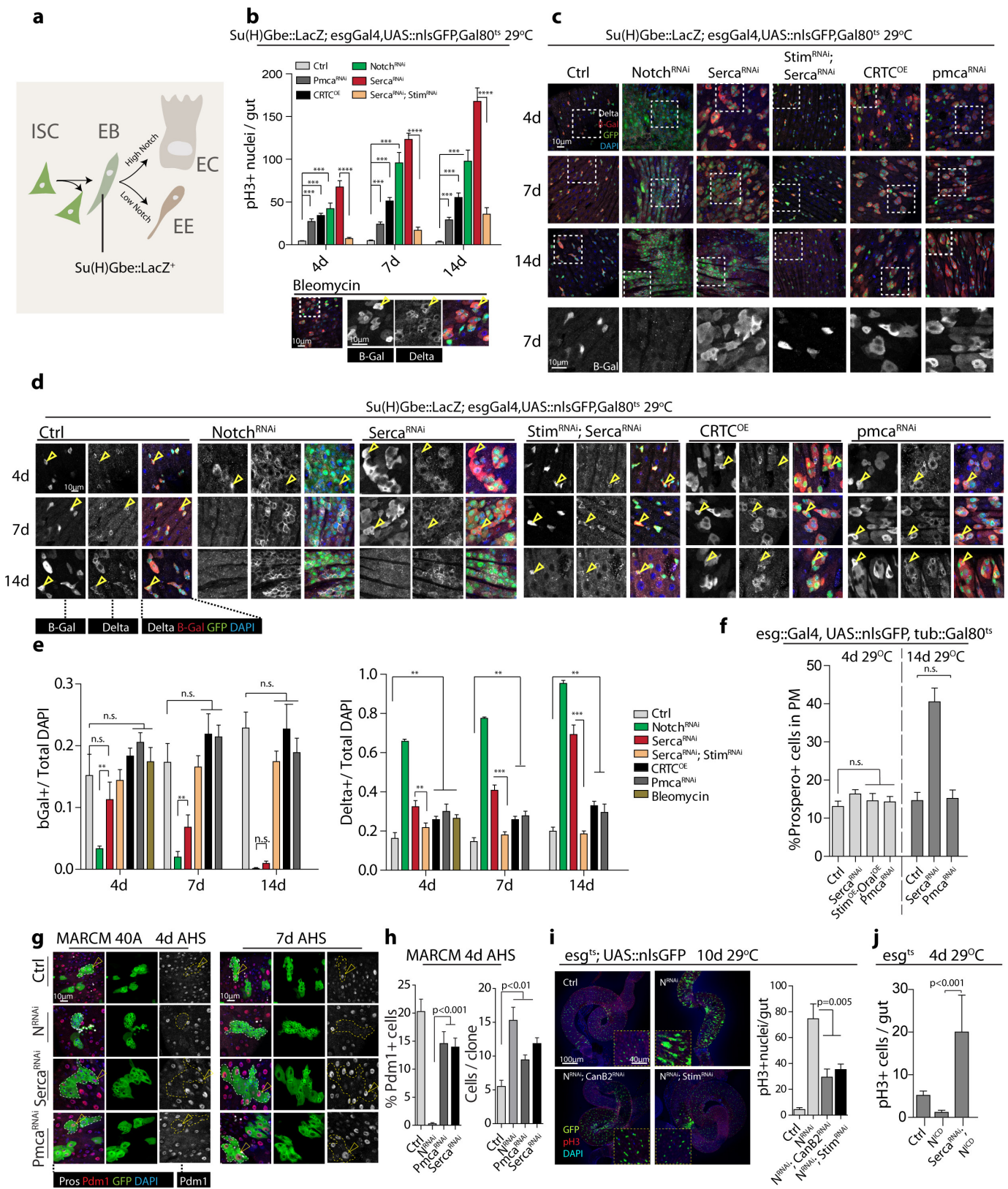




Extended Data Figure 7 | See next page for caption.

**Extended Data Figure 7 | Prolonged increase of cytosolic  $[Ca^{2+}]$  promotes ISC proliferation.** **a**, Typical traces of  $Ca^{2+}$  recordings of indicated genotypes. Transgenes were induced at 29 °C for 4 days and 3–4 guts of each genotype were recorded. Thapsigargin concentration was 2  $\mu$ M in the medium while recording. Genotypes: UAS-GCaMP3; *esg::GAL4*, UAS-mCherry; *Su(H)Gbe::GAL80*, *tub::GAL80<sup>ts</sup>* combined with *w<sup>1118</sup>* for control or with the indicated transgenes. **b**, Knockdown of SERCA in ISCs, but not enteroblasts promotes ISC proliferation. SERCA was knocked down using *esg::GAL4* (targeting ISCs and enteroblasts), *Su(H)Gbe::GAL4* (enteroblasts), or *esg::GAL4*; *Su(H)Gbe-GAL80* (ISCs) combined with *tub::GAL80<sup>ts</sup>* by incubating at 29 °C for 4 days. ISCs and/or enteroblasts are labelled by nuclear GFP (UAS-nls-GFP) in green and mitotic ISCs are stained by pH3 (red). Number of dividing ISCs (pH3<sup>+</sup>) per gut were quantified and analysed. **c**, Knocking down SERCA or PMCA promotes growth of ISC-derived clones. Typical MARCM clones expressing SERCA<sup>RNAi</sup>, PMCA<sup>RNAi</sup> and Stim<sup>RNAi</sup> 7 days after induction are shown. ISC clones are marked in green and pH3 staining was used to detect dividing ISCs. Note: SERCA<sup>RNAi</sup> clones contain many more, but smaller cells, resulting in similar clonal area compared to wild type. Quantification is shown on the right. **d**, Left, mitotic figures quantified in intestines of flies homozygous for SERCA<sup>kum170</sup>, a temperature-sensitive SERCA loss-of-function allele. Flies were exposed to heat shock (42 °C, 5 min for two consecutive days, this permanently inactivates SERCA<sup>54</sup>),

and proliferation was assessed 7 days after heat shock. Right, size of MARCM clones of ISCs homozygous for SERCA<sup>kum170</sup> analysed 7 days after clone induction. **e**, MARCM clones of ISCs homozygous for the *Stim* null allele, *Stim<sup>A</sup>*, analysed 7 days after clone induction. Dashed lines delineate individual clones (GFP, green; DAPI, blue, armadillo, red membrane; prospero, red nuclei; Df, white). Df channel is shown separately in greyscale. Quantification of clone sizes (cells per clone) is shown on the right. **f**, Guts of the indicated genotypes visualized in basal views (six panels) and in cross section (two panels). In basal views, ISCs are identified by Df immunostaining (red) and GFP (green). In cross sections, visceral muscle is identified by phalloidin staining (red), and ISCs/enteroblasts by GFP (green). Data are mean and s.e.m. *P* value from ANOVA (**c**) or Student's *t*-test (**b**, **d**, **e**). For mitotic figures in **b**, *n* = 11 for each genotype. For mitotic figures in **d** (left), *n* = 10 for each condition. For clonal analysis in **c**, clones (FRT40A, *n* = 50; SERCA<sup>RNAi</sup>, *n* = 42; Stim<sup>RNAi</sup>, *n* = 40; PMCA<sup>RNAi</sup>, *n* = 56) from 8–10 guts were quantified. For clonal analysis in **d** (right), clones (FRT42D, *n* = 80; SERCA<sup>kum170</sup>, *n* = 60) from 10 guts were quantified. For clonal analysis in **e** (FRT19A, *n* = 120; *Stim<sup>A</sup>*, *n* = 102) from 7 guts were assessed. For **a**, traces from individual ISCs from a representative gut of each indicated genotype are shown, 3–4 guts were recorded for each experiment and 2 independent experiments were performed. For **b**–**f**, 1 of 3 independent experiments is presented.



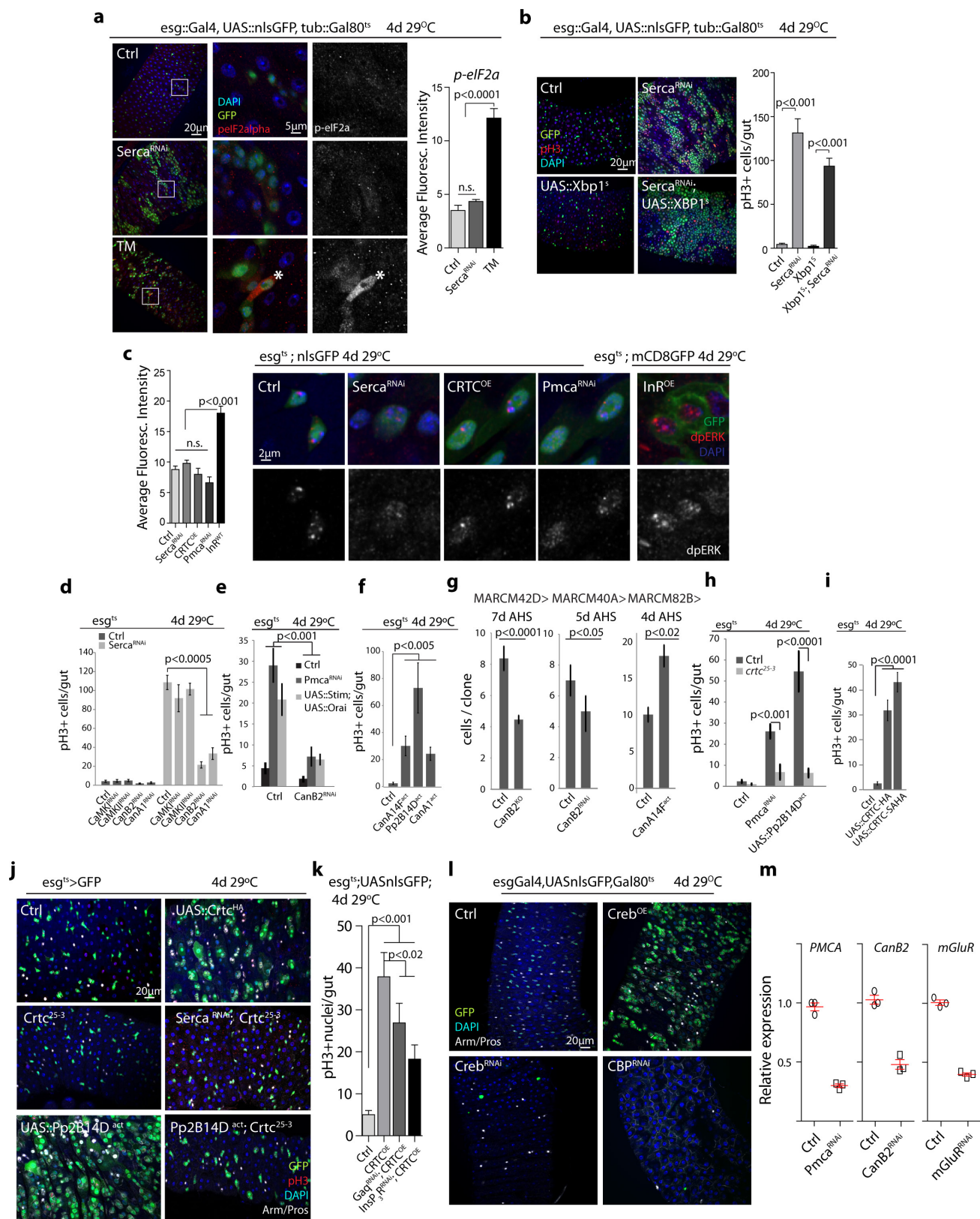
Extended Data Figure 8 | See next page for caption.



### Extended Data Figure 8 | Time-course analysis of Notch activity in ISC lineage after manipulation of $\text{Ca}^{2+}$ signalling. **a**, ISC lineage.

After an asymmetric division, ISCs (expressing Delta) activate Notch in enteroblasts (thus activating Su(H)Gbe-lacZ). Enteroblasts with high Notch activity differentiate into polyploid enterocytes (expressing Pdm1), enteroblasts with low Notch activity differentiate into enteroendocrine cells (expressing prospero). **b**, ISC proliferation rates (mitotic nuclei per gut) at different time points after perturbation of Notch (N) or  $\text{Ca}^{2+}$  signalling. Number of days after shift to  $29^{\circ}\text{C}$  is listed. **c–e**, Representative images of guts perturbed as in **b**, immunostained for Su(H)Gbe-lacZ (bGal, red, reporter for Notch activity) and D1 (white). GFP, green; DAPI, blue. For 7-day time point, higher magnification images (boxed area) of bGal channel are shown in bottom row in greyscale. Note that knockdown of Notch results in rapid loss of Su(H)Gbe-lacZ<sup>+</sup> cells and accumulation of D1<sup>+</sup> cells. SERCA knockdown results in loss of Su(H)Gbe-lacZ only after prolonged expression (14 days), even though proliferation is induced more strongly than with Notch<sup>RNAi</sup> already at 4 days. Su(H)Gbe-lacZ expression is not lost in PMCA<sup>RNAi</sup> or Crtc<sup>OE</sup> guts, although proliferation is induced as strongly as in Notch<sup>RNAi</sup>-expressing guts at 4 days. Quantification of ratio of D1<sup>+</sup> cells and Gbe-lacZ<sup>+</sup> cells in the gut after at the indicated time points after shift to  $29^{\circ}\text{C}$  is shown in **e**. **f**, Quantification of prospero<sup>+</sup> cells of posterior midguts in which cytosolic  $\text{Ca}^{2+}$  was increased by knocking down SERCA or PMCA, or by overexpressing Stim and Orai using *esg<sup>ts</sup>*. Increased numbers of enteroendocrine cells (an indication of impaired Notch signalling) were only observed in animals in which SERCA was knocked down for a prolonged period of time (14 days,  $29^{\circ}\text{C}$ ). Anti-prospero stains enteroendocrine cell nuclei. **g**, Clonal analysis of ISC differentiation process after manipulation of  $\text{Ca}^{2+}$  signalling or Notch activity. ISC MARCM clones of control, Notch<sup>RNAi</sup>, SERCA<sup>RNAi</sup> or PMCA<sup>RNAi</sup> were analysed at 4 or 7 days after heat-shock (AHS) induction. Clones (marked in GFP) are circled in dashed line. The nuclei of differentiated cells are stained by Pdm1 in red. Although PMCA<sup>RNAi</sup>

clones are significantly larger, the differentiation process is largely normal based on Pdm1 staining. Whereas differentiation of SERCA<sup>RNAi</sup> clones at 7 days after heat-shock is significantly perturbed based on Pdm1 staining, as compared with surrounding wild-type enterocytes. As expected, no Pdm1<sup>+</sup> cells were observed in Notch<sup>RNAi</sup> clones. **h**, Related to **g**, clone size and percentage of Pdm1<sup>+</sup> cells per clone was quantified 4 days after heat-shock induction. **i**, Notch<sup>RNAi</sup> induced proliferation can be partially rescued by knocking down  $\text{Ca}^{2+}$  signalling components, such as Stim or CanB2. Mitotic index shown on right. **j**, Inhibition of SERCA stimulates ISC proliferation even when Notch signalling is induced by overexpression of Notch<sup>ICD</sup> (Notch intracellular domain). Data are mean and s.e.m. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$  (ANOVA in **b**, **e**, **f**, **h** and **i**; Student's *t*-test in **b** (between SERCA<sup>RNAi</sup> and SERCA<sup>RNAi</sup>; Stim<sup>RNAi</sup>), **f** (between control and PMCA<sup>RNAi</sup>), and **j**). For **b**, mitotic cells are calculated at indicated time points for each genotype: 4 days, control ( $n = 16$ ), PMCA<sup>RNAi</sup> (8), Crtc<sup>OE</sup> (9), Notch<sup>RNAi</sup> (12), SERCA<sup>RNAi</sup> (14), SERCA<sup>RNAi</sup>; Stim<sup>RNAi</sup> (9); 7 days, control (12), PMCA<sup>RNAi</sup> (10), Crtc<sup>OE</sup> (9), Notch<sup>RNAi</sup> (11), SERCA<sup>RNAi</sup> (10), SERCA<sup>RNAi</sup>; Stim<sup>RNAi</sup> (8); 14 days, control (6), PMCA<sup>RNAi</sup> (18), Crtc<sup>OE</sup> (9), Notch<sup>RNAi</sup> (12), SERCA<sup>RNAi</sup> (8), SERCA<sup>RNAi</sup>; Stim<sup>RNAi</sup> (11). For **c–e**, cells in posterior midgut from 4–7 guts of each genotype and time point were analysed. For **e**, fraction of bGal<sup>+</sup> or D1<sup>+</sup> cells in 100–200 total cells counted in a field of the posterior midgut for each condition was quantified. Data are mean and s.e.m. for the following number of guts (left to right)  $n = 4, 4, 5, 4, 4, 4, 4, 7, 4, 6, 4, 4, 4, 5, 4, 4, 4, 5, 4, 4, 5, 4, 6, 5, 4, 4, 5, 4, 5, 5, 6, 4, 5, 7, 5, 6, 4$  and 5. For **f**, prospero<sup>+</sup> cells of indicated genotype at 4 days ( $n = 5$  guts) and 14 days ( $n = 7$  guts) were analysed. For clonal analysis in **g** and **h**, clones (FRT40A,  $n = 50$ ; SERCA<sup>RNAi</sup>,  $n = 45$ ; Notch<sup>RNAi</sup>,  $n = 30$ ; and PMCA<sup>RNAi</sup>,  $n = 43$ ) were analysed. For mitotic analysis in **i** and **j**,  $n = 12$  for each genotype. Data in **b–f** and **i** are representative of 2 independent experiments; data in **g**, **h** and **j** represent one of triplicate experiments.

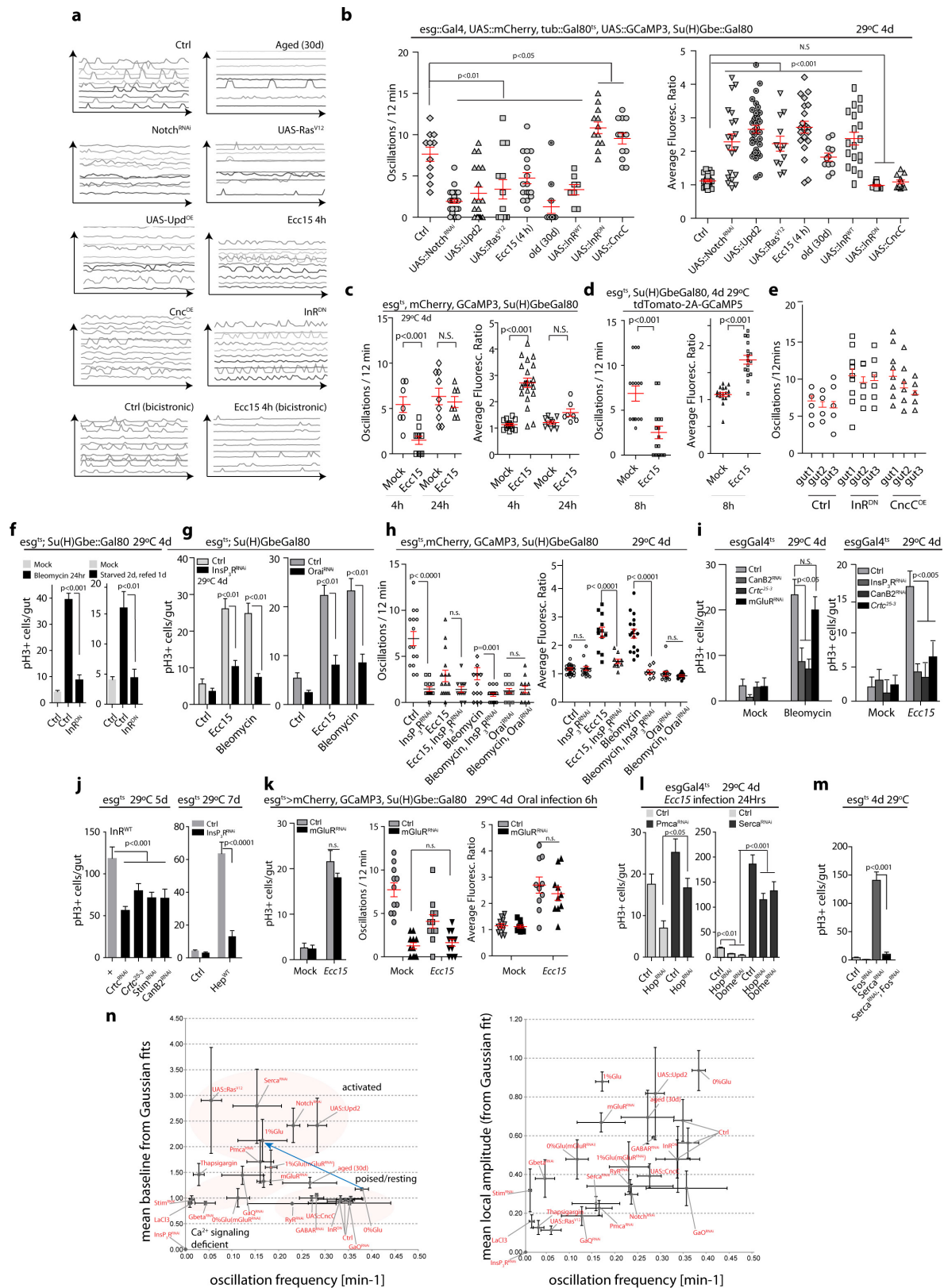


Extended Data Figure 9 | See next page for caption.

**Extended Data Figure 9 | Cytosolic  $\text{Ca}^{2+}$  regulates ISC proliferation through CaN/CRTC pathway.** **a**, Acute knockdown of *SERCA* in ISCs does not induce ER stress. Phospho-eIF2 $\alpha$  (red, a PERK-mediated phosphorylation and marker of ER stress; see also ref. 40), is only weakly detected in control ISCs. No significant increase of phospho-eIF2 $\alpha$  staining is observed in ISCs expressing *SERCA*<sup>RNAi</sup>, whereas tunicamycin treatment (an inducer of ER stress; 50  $\mu\text{M}$ , 24 h) increases phospho-eIF2 $\alpha$  strongly in ISCs (asterisk). Quantification of average fluorescent intensity of phospho-eIF2 $\alpha$  shown on right. **b**, Increasing ER folding capacity by overexpressing spliced Xbp1 (*Xbp1s*) does not limit ISC proliferation in *SERCA* loss-of-function conditions. Representative images shown on the left. Mitotic ISCs stained by anti-pH3 staining (red). Mitotic index quantified on the right. **c**, Increased cytosolic [ $\text{Ca}^{2+}$ ] induces proliferation without perturbing *Egfr* pathway activity. Indicated  $\text{Ca}^{2+}$  signalling components were knocked down or overexpressed using *esg*<sup>ts</sup>;UAS-nls-GFP. Diphospho-ERK (dpERK) staining as a readout of *Egfr* pathway activity was quantified after 4-day induction at 29 °C. InR was overexpressed using *esg*<sup>ts</sup>;UAS-mCD8-GFP as a positive control. **d**, Proliferation of *SERCA*-deficient ISCs is suppressed when CaN subunits are silenced simultaneously, but not when CaMKI and CaMKII are knocked down. **e**, Proliferation of ISCs in which cytosolic [ $\text{Ca}^{2+}$ ] is increased by overexpression of *Stmi* and *Orai*, or by knockdown of *PMCA*, is rescued when *CaNB2* is silenced simultaneously. **f**, Overexpression of constitutive active forms of CaN catalytic subunits (*CanA14F*, *Pp2B14D*, *CanA1*) promotes ISC proliferation. Mitotic figures were quantified after 4 days of transgene expression at 29 °C. **g**, CaN is required for growth of ISC lineages. Quantification of clone sizes of MARCM clones homozygous for the null allele *CanB2*<sup>KO</sup>, or expressing dsRNA against *CanB2* (*CanB2*<sup>RNAi</sup>), or constitutively active *CanA14F*<sup>act</sup>. **h**, Loss of *Crtc* (homozygosity for null allele *crtc*<sup>25-3</sup>) rescues increased ISC proliferation when *PMCA* is knocked down or when *Pp2B-14D*<sup>act</sup> is overexpressed using *esg*::*GAL4*, *tub*::*GAL80*<sup>ts</sup> (*esg*<sup>ts</sup>). **i**, Overexpression of *Crtc* promotes ISC proliferation. Mitotic index of intestines overexpressing haemagglutinin (HA)-tagged forms of wild-type *Crtc* (UAS-CRTC-HA) or constitutively nuclear forms of *Crtc* (UAS-CRTC-SA-HA) using *esg*::*GAL4*<sup>ts</sup> was analysed after 4 days of incubation at 29 °C. **j**, Increasing cytosolic  $\text{Ca}^{2+}$  promotes ISC proliferation via the  $\text{Ca}^{2+}$ /calcineurin/*Crtc* pathway. Guts of indicated genotypes were dissected and stained with anti-pH3 (indicating mitotic ISCs), anti-armadillo (arm, labelling cell boundaries) and anti-prospero

(pro, labelling enteroendocrine cells). **k**, *Crtc* overexpression is sufficient to promote ISC proliferation when *G $\alpha$ q* or *InsP3R* are silenced. **l**, CREB and its partner CBP are required for ISC survival, while overexpressing CREB promotes proliferation. Representative images of intestines in which CREB and CBP were genetically perturbed in ISC/enteroblasts using *esg*::*GAL4*. Increased numbers of GFP<sup>+</sup> ISCs/enteroblasts are observed when CREB is overexpressed, while a significant loss of GFP<sup>+</sup> cells is observed when CREB or CBP are knocked down. Guts are stained for armadillo (membrane, white) and prospero (nuclear, white) to identify enteroendocrine cells and enterocytes; DAPI, blue. Genotype: *esg*::*GAL4*; UAS-nls-GFP; *tub*::*GAL80*<sup>ts</sup>/UAS-X. **m**, Knockdown efficiency of RNAi lines determined by qRT-PCR. *PMCA*<sup>RNAi</sup> (BL31572) and *CanB2*<sup>RNAi</sup> (BL27270) were used to knock down respective genes in the gut using NP1::*GAL4*, *tub*::*GAL80*<sup>ts</sup> (29 °C for 10 days). *mGluR*<sup>RNAi</sup> (BL41668) was used to knock down *mGluR* in the brain using *elav*::*GAL4*. Expression levels were normalized using *Act5C* and to uninduced controls (NP1::*GAL4*/+; *tub*::*GAL80*<sup>ts</sup>/+ or *elav*::*GAL4*/+). Effectiveness of other constructs used has been reported in the literature: UAS-*G $\alpha$ q*<sup>RNAi</sup> and UAS-*PLC $\beta$* <sup>RNAi</sup> were obtained from and verified by Ha *et al.*<sup>55</sup>, UAS-*CamKI*<sup>RNAi</sup> (BL26726), UAS-*CamKII*<sup>RNAi</sup> (BL29401), UAS-*InsP3R*<sup>RNAi</sup> (BL25937), and UAS-*RyR*<sup>RNAi</sup> (BL28919) were verified by Shim *et al.*<sup>23</sup>, and UAS-*SERCA*<sup>RNAi</sup> (BL44581) by Roti *et al.*<sup>29</sup>. Data are mean and s.e.m. *P* values from Student's *t*-test (**b**, **g** and **h**) and ANOVA (**a**, **c**, **d**–**f**, **i** and **k**). For **a** and **c**, fluorescence intensities for phospho-eIF2 $\alpha$  (**a**) or diphospho-ERK (**c**) in 30–50 ISCs/enteroblasts doublets in single fields of several independent posterior midguts were averaged for each condition. In **a**, control, *n* = 5 guts, *SERCA*<sup>RNAi</sup>, *n* = 5, tunicamycin, *n* = 6; in **c**, control, *n* = 6, *SERCA*<sup>RNAi</sup>, *n* = 4, *CRTC*<sup>OE</sup>, *n* = 9, and *PMCA*<sup>RNAi</sup>, *n* = 8. For mitotic analysis, in: **b**, *n* = 13 guts for control, *n* = 10 for the rest; **d**, *n* = 17 for each genotype; **e**, *n* = 18 for control, *n* = 12 for the rest; **f** and **k**, *n* = 11 each genotype; and **i**, *n* = 12 for each condition. For clonal analysis in **g**, clones (FRT42D, *n* = 56 clones; *CanB2*<sup>KO</sup>, *n* = 60; FRT40A, *n* = 50; *CanB2*<sup>RNAi</sup>, *n* = 70; FRT82B, *n* = 62; *CanA14F*<sup>act</sup>, *n* = 58) from 10 guts each were assessed. One representative image from 10 flies in a single experiment (two independent experiments) is shown in **j** and **l**. Data in **a**–**i**, and **k** are representative of three independent experiments. Data in **m** is from *n* = 3 technical replicates of samples pooled from 10 guts each for *PMCA* and *CanB2*, 4 heads each for *mGluR*. Representative of two independent experiments.





Extended Data Figure 10 | See next page for caption.

**Extended Data Figure 10 |  $\text{Ca}^{2+}$  oscillation pattern as an indicator of ISC proliferation status.**

**a**, Typical traces of live recordings of indicated genotypes. Genotype for GCaMP3 control:  $w^{1118}$  X UAS-GCaMP3;  $\text{esg}::\text{GAL4}$ , UAS-mCherry; Su(H)Gbe::GAL80, tub::GAL80<sup>ts</sup>. Genotype for bicistronic control:  $w^{1118}$  X;  $\text{esg}::\text{GAL4}$ , UAS-tdTomato-P2A-GCaMP5; Su(H)Gbe::GAL80, tub::GAL80<sup>ts</sup>. **b**,  $\text{Ca}^{2+}$  oscillation patterns of ISCs in which proliferation was stimulated or inhibited by genetic or environmental perturbations: knockdown of Notch, overexpression of InR, Unpaired2 or Ras<sup>V12</sup>, infection with *Ecc15*, or ageing results in high proliferative activity. Overexpression of InR<sup>DN</sup> and CncC inhibits proliferation of ISCs. **c**, Acute *Ecc15* infection transiently increases cytosolic  $[\text{Ca}^{2+}]$  while decreasing oscillations in ISCs. **d**, Acute *Ecc15* infection increases cytosolic  $\text{Ca}^{2+}$  while decreasing oscillations in ISCs as determined using the bi-cistronic calcium reporter UAS-tdTomato-P2A-GCaMP5G. **e**, ISCs in which proliferation is impaired exhibit more frequent oscillations than controls. Oscillation frequency of ISCs from indicated genotypes is plotted individually (three guts for each genotype and each dot represents one ISC). **f**, Expressing InR<sup>DN</sup> in ISCs is sufficient to inhibit stress- or diet-induced proliferation. Quantification of mitotic figures of indicated genotype is shown. For bleomycin treatment, flies were dry-starved for 4 h before feeding on  $25 \mu\text{g ml}^{-1}$  (final) bleomycin for 24 h. For refeeding, flies were maintained on normal food for 4 days at 29 °C, then starved for 2 days, and re-fed with yeast-supplemented food. Genotype:  $\text{esg}::\text{GAL4}$ , UAS-GFP; Su(H)Gbe::GAL80, tub::GAL80<sup>ts</sup>/UAS-InR<sup>DN</sup>. **g**, ISC proliferation induced by oral infection with *Ecc15* or by bleomycin treatment is suppressed by silencing *InsP3R* or *Orai*. Mitotic figures were analysed 6 h after oral infection with *Ecc15* or 24 h after feeding with bleomycin. **h**, Increased cytosolic  $[\text{Ca}^{2+}]$  in ISCs activated by oral infection with *Ecc15* or by bleomycin treatment. This increase is suppressed by silencing *InsP3R* or *Orai*. **i**, ISC proliferation induced by bleomycin treatment or *Ecc15* infection is suppressed by silencing *Crtc* or *InsP3R* or *CanB2*, or in *crtc*<sup>25-3</sup> homozygous mutants. *mGluR*, in turn, is not required for bleomycin-induced proliferation. Mitotic figures were analysed 24 h after feeding with bleomycin or 6 h after oral infection with *Ecc15*. **j**, Quantification of mitotic figures in animals of the indicated genotypes. ISC proliferation

induced by overexpression of InR or Hep can be suppressed by knockdown of *InsP3R*, *Crtc*, *Stim* or *CanB2* (see also Fig. 4a). **k**, *mGluR* is not required for *Ecc15*-induced proliferation and changes in cytosolic  $\text{Ca}^{2+}$ . Mitotic figures and  $\text{Ca}^{2+}$  oscillation patterns analysed 6 h after oral infection. **l**, Increasing cytosolic  $[\text{Ca}^{2+}]$  promotes ISCs proliferation in Jak/Stat loss-of-function conditions. Jak/Stat pathway (*Dome* and *Hop*) is required for ISC proliferation induced by *Ecc15* infection. Increasing cytosolic  $[\text{Ca}^{2+}]$  by knocking down *PMCA* (left) or *SERCA* (right) is sufficient to rescue ISC proliferation. **m**, Knocking down *Fos* can substantially suppress *Crtc* overexpression induced ISC proliferation. **n**, Left, segregation of active and resting ISCs into  $\text{Ca}^{2+}$  oscillation modes as calculated by automatic peak detection and Gaussian fits. As shown for 'manual' calculations in Fig. 4,  $\text{Ca}^{2+}$  oscillation patterns segregate into two modes associated with the proliferative status of the ISCs. ISCs in which the core components of  $\text{Ca}^{2+}$  homeostatic machinery are perturbed exhibit both low oscillation frequency and low average signal intensity (bottom left corner). Transition from quiescence to active proliferation after L-Glu feeding is indicated by the blue arrow. Right, ISC activity does not segregate when local oscillation amplitudes are plotted against oscillation frequency, suggesting that the primary driver of ISC proliferation is not the amplitude of individual  $\text{Ca}^{2+}$  spikes, but the increase in basal or average cytosolic  $\text{Ca}^{2+}$  concentration within ISC. Data are mean and s.e.m. *P* values from ANOVA (**b**, **h**, **i**, **j** (left) and **l** (right)) and Student's *t*-test (**c**, **d**, **f**, **g**, **k**, **m**, **j** (right) and **l** (left)). For  $\text{Ca}^{2+}$  recordings in **b**–**e**, **h** and **k**, individual ISCs pooled from 3–4 guts were plotted. The sample size for **b** is  $n = 10, 14, 14, 12, 16, 18, 9, 12, 11, 20, 22, 38, 29, 12, 19, 8$  and 12; for **c**,  $n = 7, 8, 9, 7, 13, 18, 10$  and 7; for **d**,  $n = 14, 16, 15$  and 18; for **e**,  $n = 5, 4, 5, 9, 4, 6, 8, 6$  and 5; for **h**,  $n = 12, 11, 12, 11, 12, 14, 12, 11, 17, 12, 14, 11, 19, 9, 10$  and 8; for **k**,  $n = 11, 12, 11, 11, 13, 10, 11$  and 11 (from left to right for all panels). For mitotic analysis in **f** and **g**,  $n = 12, 10, 10, 10, 12, 11, 10, 12, 12, 14, 10, 9, 9, 10, 9, 10$  and 13; and in **i** and **m**,  $n = 13, 12, 11, 18, 11, 18, 12, 13, 13, 12, 10, 10, 13, 9, 12, 12, 16, 12, 9, 10, 10, 8, 9, 9, 10, 9, 12, 13, 13, 14, 14, 15, 12, 10, 19, 10, 12, 10, 9, 8, 8, 9$  and 10 (from left to right for all panels). Data in **a**–**h**, **j** and **m** are representative of three independently performed experiments, and those shown in **i**, **k** and **l** are a composite from two separate experiments.

# The histone chaperone CAF-1 safeguards somatic cell identity

Sihe Cheloufi<sup>1,2,3\*</sup>, Ulrich Elling<sup>4\*</sup>, Barbara Hopfgartner<sup>5</sup>, Youngsook L. Jung<sup>6,7</sup>, Jernej Murn<sup>8,9</sup>, Maria Ninova<sup>10</sup>, Maria Hubmann<sup>4</sup>, Aimee I. Badeaux<sup>8,9</sup>, Cheen Euong Ang<sup>11</sup>, Danielle Tenen<sup>2,12</sup>, Daniel J. Wesche<sup>1,2,3</sup>, Nadezhda Abazova<sup>1,2,3</sup>, Max Hogue<sup>1,2,3</sup>, Nilgun Tasdemir<sup>13</sup>, Justin Brumbaugh<sup>1,2,3</sup>, Philipp Rathert<sup>5</sup>, Julian Jude<sup>5</sup>, Francesco Ferrari<sup>6,7</sup>, Andres Blanco<sup>8,9</sup>, Michaela Fellner<sup>5</sup>, Daniel Wenzel<sup>4</sup>, Marietta Zinner<sup>4</sup>, Simon E. Vidal<sup>14</sup>, Oliver Bell<sup>4</sup>, Matthias Stadtfeld<sup>14</sup>, Howard Y. Chang<sup>3,15</sup>, Genevieve Almouzni<sup>16</sup>, Scott W. Lowe<sup>3,13</sup>, John Rinn<sup>2,12</sup>, Marius Wernig<sup>11</sup>, Alexei Aravin<sup>10</sup>, Yang Shi<sup>8,9</sup>, Peter J. Park<sup>6,7</sup>, Josef M. Penninger<sup>4</sup>, Johannes Zuber<sup>5§</sup> & Konrad Hochedlinger<sup>1,2,3§</sup>

**Cellular differentiation involves profound remodelling of chromatic landscapes, yet the mechanisms by which somatic cell identity is subsequently maintained remain incompletely understood. To further elucidate regulatory pathways that safeguard the somatic state, we performed two comprehensive RNA interference (RNAi) screens targeting chromatin factors during transcription-factor-mediated reprogramming of mouse fibroblasts to induced pluripotent stem cells (iPS cells). Subunits of the chromatin assembly factor-1 (CAF-1) complex, including Chaf1a and Chaf1b, emerged as the most prominent hits from both screens, followed by modulators of lysine sumoylation and heterochromatin maintenance. Optimal modulation of both CAF-1 and transcription factor levels increased reprogramming efficiency by several orders of magnitude and facilitated iPS cell formation in as little as 4 days. Mechanistically, CAF-1 suppression led to a more accessible chromatin structure at enhancer elements early during reprogramming. These changes were accompanied by a decrease in somatic heterochromatin domains, increased binding of Sox2 to pluripotency-specific targets and activation of associated genes. Notably, suppression of CAF-1 also enhanced the direct conversion of B cells into macrophages and fibroblasts into neurons. Together, our findings reveal the histone chaperone CAF-1 to be a novel regulator of somatic cell identity during transcription-factor-induced cell-fate transitions and provide a potential strategy to modulate cellular plasticity in a regenerative setting.**

Ectopic expression of transcription factors is sufficient to override stable epigenetic programs and hence alter cell fate<sup>1</sup>. For example, forced expression of the pluripotency-related transcription factors Oct4, Klf4, Sox2 and c-Myc (OKSM) in somatic cells yields iPS cells, which are molecularly and functionally equivalent to embryonic stem cells (ES cells)<sup>2</sup>. Similarly, ectopic expression of lineage-specific transcription factors drives conversion of heterologous cells into cardiac, neuronal, myeloid and other specialized cell types<sup>3</sup>. However, the reprogramming process is generally slow and inefficient, suggesting that chromatin-associated mechanisms are in place to safeguard somatic cell identity and confer resistance to cell-fate change.

Previous efforts to identify chromatin modulators of iPS cell formation included gain- and loss-of-function screens, as well as transcriptional profiling of bulk or FACS-enriched cell populations undergoing reprogramming. However, iPS cell modulators that do not change transcriptionally are typically overlooked when analysing expression dynamics in reprogramming intermediates<sup>4</sup>. Moreover, known repressors of iPS cell formation such as p53, Mbd3, Dot1l and Dnmt1 were either predicted or identified from small candidate sets and some of these molecules appear to depend on specific cell contexts or culture conditions<sup>5–7</sup>. Although

large-scale RNAi screens have been used to systematically probe roadblocks to reprogramming<sup>4,8,9</sup>, this approach remains technically challenging due to the lack of effective short hairpin RNAs (shRNAs), prevalent off-target effects, and biases in the library representation or the screening readout. We therefore hypothesized that additional barriers to iPS cell formation remain to be discovered and should yield insights into mechanisms that safeguard somatic cell identity.

To systematically explore chromatin factors that resist transcription-factor-induced cell-fate transitions, we used custom microRNA-based shRNA libraries targeting known and predicted chromatin regulators in two independent screening strategies during the reprogramming of fibroblasts into iPS cells. Both screens validated previously implicated chromatin pathways and revealed novel, potent repressors of reprogramming. Through a series of cellular and molecular studies, we found that suppression of a histone chaperone complex markedly enhanced and accelerated iPS cell formation by influencing local chromatin accessibility, transcription factor binding and histone H3K9 trimethylation (H3K9me3). We propose that this complex functions as a key determinant of cellular identity by resisting transcription-factor-induced cell-fate change.

<sup>1</sup>Department of Molecular Biology, Cancer Center and Center for Regenerative Medicine, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. <sup>2</sup>Department of Stem Cell and Regenerative Biology and Harvard Stem Cell Institute, Cambridge, Massachusetts 02138, USA. <sup>3</sup>Howard Hughes Medical Institute, Chevy Chase, Maryland 20815, USA. <sup>4</sup>Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), Vienna Biocenter (VBC), A-1030 Vienna, Austria. <sup>5</sup>Research Institute of Molecular Pathology (IMP), Vienna Biocenter (VBC), A-1030 Vienna, Austria. <sup>6</sup>Department of Biomedical Informatics, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>7</sup>Division of Genetics, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. <sup>8</sup>Department of Cell Biology, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>9</sup>Division of Newborn Medicine, Boston Children's Hospital, Boston, Massachusetts 02115, USA. <sup>10</sup>California Institute of Technology, Division of Biology and Biological Engineering, Pasadena, California 91125, USA. <sup>11</sup>Institute for Stem Cell Biology and Regenerative Medicine, Department of Pathology and Department of Bioengineering, Stanford University, Stanford, California 94305, USA. <sup>12</sup>Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, Massachusetts 02142, USA. <sup>13</sup>Memorial Sloan Kettering Cancer Center, New York, New York 10065, USA. <sup>14</sup>The Helen L. and Martin S. Kimmel Center for Biology and Medicine, Skirball Institute of Biomolecular Medicine, Department of Cell Biology, NYU School of Medicine, New York, New York 10016, USA. <sup>15</sup>Center for Personal Dynamic Regulomes and Program in Epithelial Biology, Stanford University School of Medicine, Stanford, California 94305, USA. <sup>16</sup>Centre de Recherche, Institut Curie, 75248 Paris, France.

\*These authors contributed equally to this work.

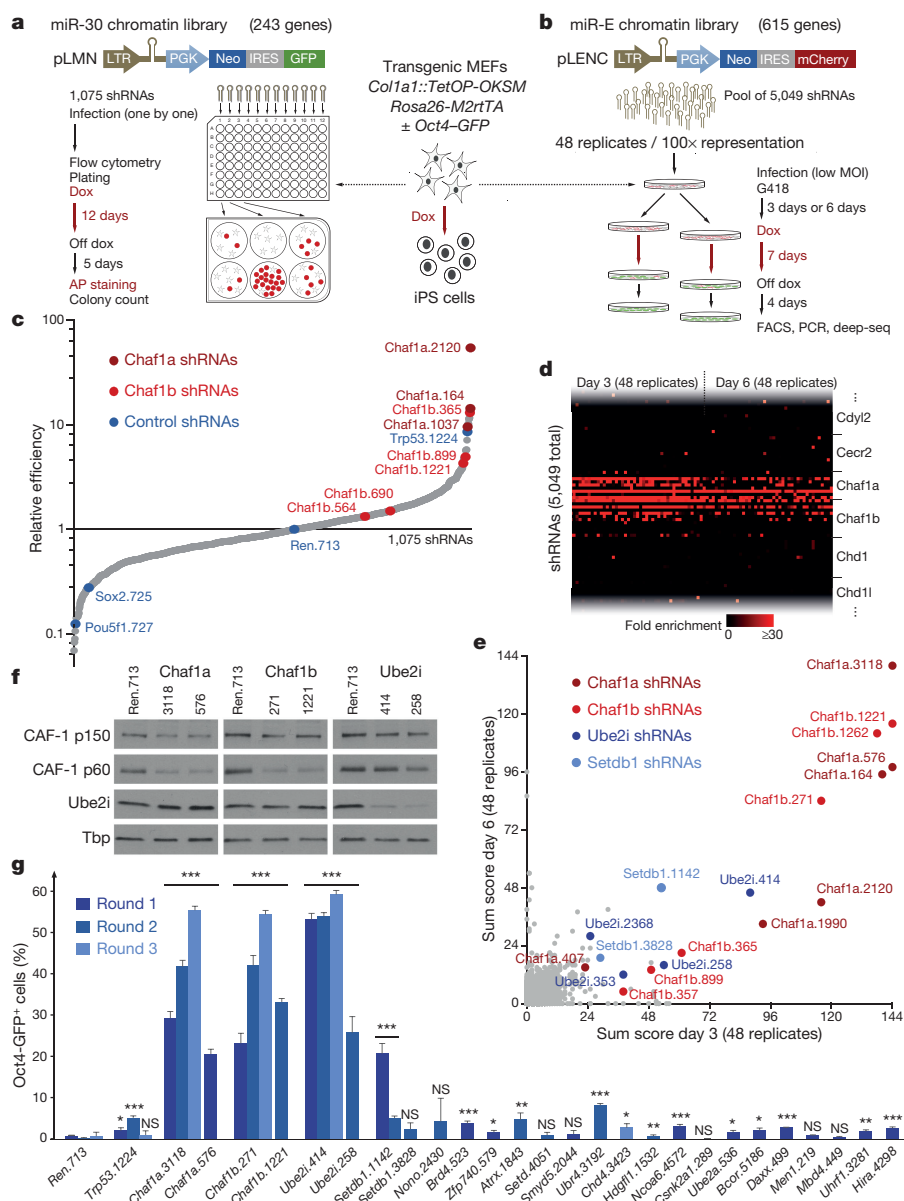
§These authors jointly supervised this work.



## RNAi screens for chromatin barriers to reprogramming

We conceived two parallel strategies for screening chromatin-focused microRNA-based shRNA (shRNAmiR) libraries in transgenic ('reprogrammable') mouse embryonic fibroblasts (MEF) harbouring a doxycycline-inducible polycistronic OKSM cassette and a constitutive M2rtTA driver<sup>10</sup>. We first designed an arrayed screening strategy using a previously described miR-30-based retroviral shRNA library targeting 243 genes<sup>11</sup> (1,071 shRNAmiRs in pLMN vector) introduced one by one into reprogrammable MEFs (Fig. 1a and Supplementary Table 1). Alkaline phosphatase-positive (AP<sup>+</sup>), transgene-independent iPS-cell-like colonies were quantified using customized image analysis software after 12 days of doxycycline exposure and 5 days of doxycycline-independent growth. Reprogramming efficiency ratios were calculated relative to a control shRNA targeting *Renilla* luciferase (Ren.713 (see Methods for details)).

In an independent multiplexed screen, we introduced an optimized miR-E-based<sup>12</sup> retroviral library targeting 615 known and predicted chromatin regulators (5,049 shRNAmiRs in pLENC) (Fig. 1b and Supplementary Table 2) into reprogrammable MEFs harbouring an Oct4-GFP reporter<sup>10</sup>. To control for biases due to background reprogramming events, we conducted parallel multiplex screens in a large number of biological replicates. Specifically, we transduced reprogrammable MEFs with the entire pool of 5,049 shRNAs in 48 biological replicates (>100 infected cells per shRNA and replicate) and induced OKSM expression 3 and 6 days after viral transduction to control for differences in target protein half-life, yielding 96 replicates in total (Fig. 1b). Library representation was then quantified by deep sequencing of transgene-independent Oct4-GFP<sup>+</sup> iPS cells isolated by FACS, and strong shRNAmiRs were identified using an additive score reflecting the consistency of shRNA enrichment across all replicates.



**Figure 1 | Arrayed and multiplexed shRNAmiR screening strategies to identify suppressors of reprogramming.** **a**, **b**, Schematic of arrayed (**a**) and multiplexed (**b**) RNAi screens. Dox, doxycycline. **c**, Results from arrayed screen, depicting average reprogramming efficiency ratios of two biological replicates normalized to *Renilla* (Ren.713) shRNA control. **d**, Heatmap depicting enrichment of selected shRNAs (shown in rows, ordered by gene symbol) over all 96 replicates (columns). **e**, Scatter plot

representing sum score of enriched shRNAs across all replicates. **f**, Western blot analysis confirming shRNA suppression of CAF-1 p150 (Chaf1a), CAF-1 p60 (Chaf1b) and Ube2i at day 3 of reprogramming (see Supplementary Fig. 1 for full scans). **g**, Validation of hits from multiplex screen. Values are the mean from biological triplicates; error bars indicate standard deviation (\* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ).

## Nucleosome assembly factor blocks iPS cell formation

Remarkably, despite methodological differences, the most prominent hits that emerged from both screens were Chaf1a and Chaf1b, two subunits of the chromatin assembly factor complex CAF-1 involved in the deposition of canonical histones H3/H4 on newly synthesized DNA<sup>13</sup> (Fig. 1c–e and Supplementary Tables 1 and 2). Additional hits of interest included the SUMO-conjugating enzyme Ube2i, a novel repressor of iPS cell formation, as well as the SET domain-containing H3K9 methyltransferase Setdb1, which has previously been shown to inhibit reprogramming<sup>14–16</sup>. Importantly, the top-scoring shRNAs targeting Chaf1a, Chaf1b, or Ube2i reduced the expression of their predicted target genes (Fig. 1f, Extended Data Fig. 1a–d and Supplementary Fig. 1).

We validated enhanced reprogramming with top-scoring shRNAs from the multiplexed screen (Fig. 1g and Supplementary Table 2). Although reprogramming efficiency increased from 0.5% in controls to 1–5% with Trp53 shRNAs, consistent with previous reports<sup>17</sup>, Chaf1a, Chaf1b, or Ube2i suppression robustly enhanced reprogramming efficiency up to 20–60%, regardless of culture conditions (Fig. 1g and Extended Data Fig. 1e, f). shRNAs targeting Sae1 and Uba2a, which are components of the SUMO E1 ligase complex, mirrored the effects of Ube2i suppression on reprogramming (Extended Data Fig. 1g), further validating the sumoylation pathway as a major repressor of iPS cell formation.

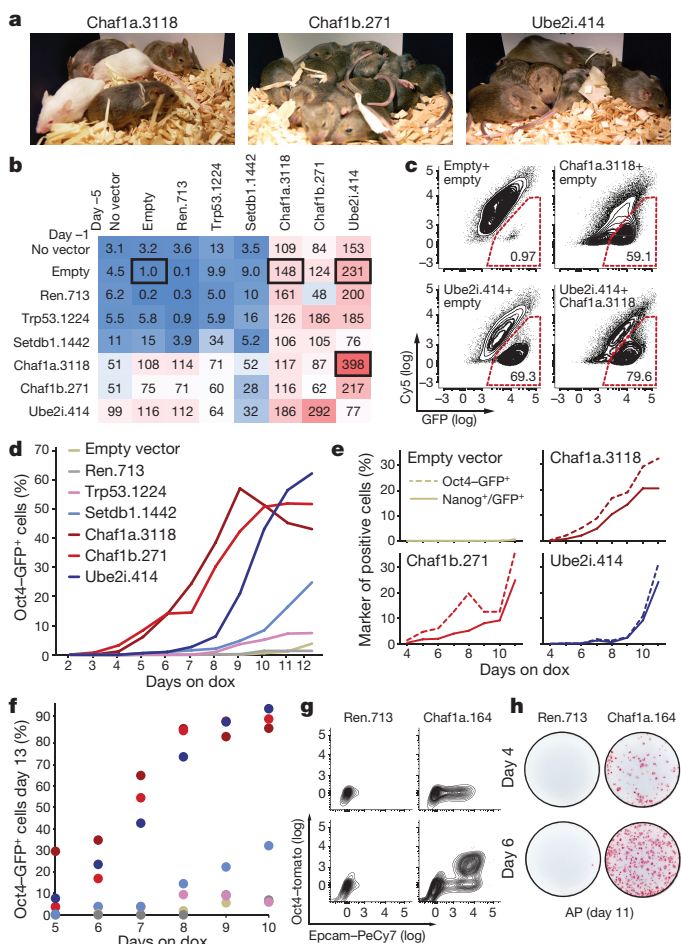
We next used a doxycycline-inducible shRNAmiR expression system to test whether transient suppression of Chaf1a, Chaf1b or Ube2i during reprogramming affected the ability of the resultant iPS cells to contribute to normal development. shRNAs and OKSM were simultaneously expressed in reprogrammable MEFs for 7 days. Oct4-GFP<sup>+</sup> cells were then purified using FACS, followed by doxycycline withdrawal to select for transgene-independent iPS cells. Injection of these iPS cells into blastocysts gave rise to adult high-grade chimaeras that efficiently produced germline offspring (Fig. 2a and Extended Data Fig. 2a). These results indicate that silencing of Chaf1a, Chaf1b or Ube2i during reprogramming does not compromise the potential of iPS cells to differentiate into somatic and germ cell lineages *in vivo*.

## Functional interactions of CAF-1, Ube2i and Setdb1

We next devised a combinatorial RNAi approach to assess whether the reprogramming barriers Chaf1a, Chaf1b, Ube2i and Setdb1 interact during iPS cell generation. Constitutive miR-E-based shRNA vectors targeting different pairs of chromatin regulators were sequentially transduced into reprogrammable MEFs, which were subsequently induced with doxycycline for 7 days. The transgene-independent fraction of reprogrammed cells at day 11 was then determined by flow cytometry and plotted as the ratio of Oct4-GFP<sup>+</sup> to Oct4-GFP<sup>-</sup> cells relative to an empty vector control (Fig. 2b, c and Extended Data Fig. 2b). Co-suppression of two CAF-1 subunits, or of Setdb1 plus Chaf1a, Chaf1b or Ube2i impaired overall reprogramming efficiencies compared to suppression of each gene alone, although we cannot exclude that this phenotype is due to cellular toxicity elicited by potent suppression of these chromatin regulators. By contrast, simultaneous suppression of either CAF-1 subunit and Ube2i further increased reprogramming efficiency, suggesting that these factors may act in independent pathways or stages to repress iPS cell formation.

## Suppression of CAF-1 accelerates iPS cell formation

To gain insights into the dynamics of reprogramming in the absence of the identified chromatin barriers, we followed the emergence of Oct4-GFP<sup>+</sup> cells over time. Although suppression of Ube2i promoted Oct4-GFP activation slightly earlier than controls (day 6 with Ube2i shRNA versus day 9 with *Renilla* shRNA), the suppression of either CAF-1 subunit triggered a dramatic acceleration of this process and consistently generated Oct4-GFP<sup>+</sup> cells as early as 5 days after OKSM expression (Fig. 2d). Similar effects were observed with analysis of Nanog expression (Fig. 2e and Extended Data Fig. 2c).



**Figure 2 | CAF-1 suppression accelerates reprogramming and yields developmentally competent iPS cells.** **a**, Generation of iPS-cell-derived chimaeras using indicated shRNAmiRs. **b**, Combinatorial RNAi studies involving pairs of shRNAmiRs sequentially transduced 5 and 1 day before OKSM induction. Reprogramming efficiencies are shown as ratio of Oct4-GFP<sup>+</sup> to Oct4-GFP<sup>-</sup> cells at day 11 relative to an empty vector control. **c**, Flow cytometry plots of representative samples used for **b**. **d**, **e**, Time-course analysis of Oct4-GFP (**d**) and Nanog (**e**) expression during reprogramming of MEFs transduced with indicated shRNAmiRs. **f**, Establishment of transgene-independent Oct4-GFP<sup>+</sup> iPS cells in the presence of indicated shRNAmiRs (colour code as in **d**). Samples were induced with doxycycline for indicated number of days and analysed at day 13. **g**, Expression dynamics of reprogramming markers Epcam and Oct4-tomato after 4 and 6 days of OKSM expression (media supplemented with 2i, ascorbate and Dot11 inhibitor). **h**, Alkaline phosphatase (AP)-positive iPS cell colonies scored at day 11 after 4 or 6 days of OKSM expression (representative example from two biological replicates and three technical replicates).

We next examined the ability of our candidate shRNAs to facilitate transgene-independent clonal growth, a hallmark of iPS cells. Suppression of either CAF-1 subunit or Ube2i gave rise to transgene-independent Oct4-GFP<sup>+</sup> cells after 5 days of OKSM expression, the earliest examined time point, whereas stable iPS cells only appeared by day 9 in control shRNA-treated cells (Fig. 2f). Examination of cell surface markers associated with different stages of reprogramming further showed that a larger fraction of Chaf1a shRNA vector-infected cells expressed the early intermediate marker Epcam at day 4 of iPS cell formation compared to controls, with a subset of cells activating the Oct4-tomato reporter (Fig. 2g). Consistently, inhibition of Chaf1a consistently supported the formation of AP<sup>+</sup>, transgene-independent iPS cells after as little as 4 days of factor expression (Fig. 2h). Based on the notable effects of Chaf1a and Chaf1b suppression on the dynamics and efficiency of iPS cell formation, we focused subsequent analyses on these two components of the CAF-1 complex.

Suppression of CAF-1 subunits had no major effects on OKSM expression at the RNA or protein level, indicating that the observed phenotype is due to direct modulation of the reprogramming transgenes (Extended Data Fig. 3a–c and Supplementary Fig. 1). Moreover, the reprogramming increase elicited by Chaf1b shRNAs could be rescued by overexpression of an shRNA-resistant version of human CHAF1B cDNA, demonstrating specificity of the effect (Extended Data Fig. 3d). Lastly, knockdown of either Chaf1a or Chaf1b did not increase cell proliferation in the presence of OKSM induction, indicating a growth-independent effect of CAF-1 suppression on reprogramming efficiency (Extended Data Fig. 3e, f).

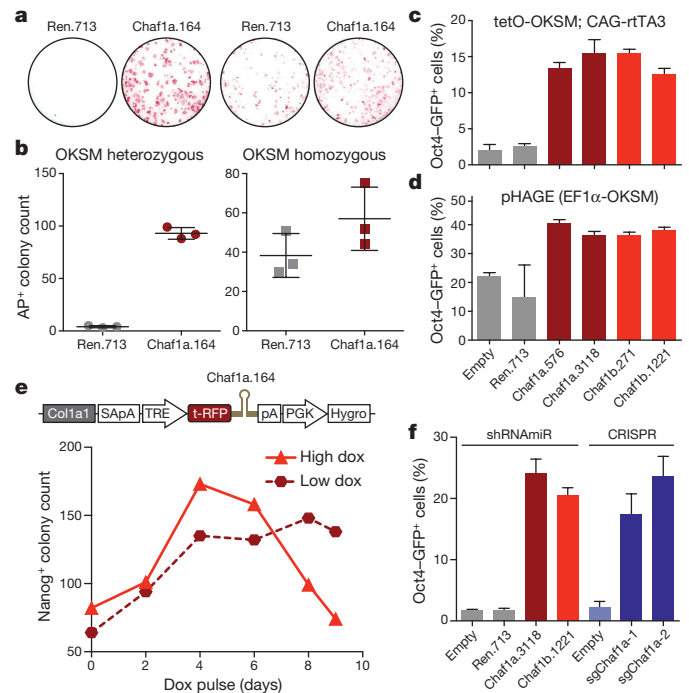
### Enhanced reprogramming requires optimal CAF-1 dosage

To determine whether the effect of CAF-1 on reprogramming depends on OKSM expression levels, we compared iPS cell formation between reprogrammable MEFs carrying either one (heterozygous) or two (homozygous) copies of the *Col1a1::tetOP-OKSM* and *R26-M2rtTA* alleles<sup>10,18</sup>. Although CAF-1 suppression in heterozygous MEFs enhanced iPS cell formation by orders of magnitude, CAF-1 suppression in homozygous MEFs resulted in a more modest increase in iPS cell numbers (Fig. 3a, b). Accordingly, we observed that CAF-1 knockdown had a stronger effect on iPS cell derivation efficiency when infecting MEFs with viral vectors achieving moderate or transient OKSM expression compared to vectors achieving high OKSM expression (Fig. 3c, d and Extended Fig. 4a–e). These results show that the reprogramming phenotype induced by CAF-1 suppression is influenced by both the levels and the duration of OKSM expression.

CAF-1 is essential for embryonic growth and viability of cultured cells in the absence of exogenous OKSM expression<sup>19–22</sup>, which we confirmed in NIH3T3 cells using the strongest CAF-1 shRNAs (Extended Data Fig. 5a). To test whether the duration and degree of CAF-1 suppression might affect reprogramming efficiency, we generated transgenic MEFs carrying a doxycycline-inducible Chaf1a shRNA linked to an RFP reporter in the *Col1a1* locus (Fig. 3e). Infection of transgenic MEFs with a constitutive lentiviral vector expressing OKSM in the presence of low doses of doxycycline ( $0.2 \mu\text{g ml}^{-1}$ ) for 2–9 days resulted in a progressive increase in the formation of Nanog<sup>+</sup> iPS cell colonies, which plateaued by day 6 (Fig. 3e and Extended Data Fig. 5b–d). By contrast, exposure of replicate cultures to high doses of doxycycline ( $2 \mu\text{g ml}^{-1}$ ) increased reprogramming efficiency until day 4 but decreased iPS cell colony numbers thereafter. These data suggest that enhanced reprogramming is also dependent on CAF-1 dose, with early CAF-1 suppression being beneficial but long-term, potent suppression being detrimental to iPS cell derivation.

To investigate whether genomic perturbation of CAF-1 components mimics the phenotype elicited by shRNAs, we introduced mutations into the endogenous *Chaf1a* locus using CRISPR/Cas9 technology. Briefly, reprogrammable MEFs were transduced with lentiviral vectors expressing Cas9 and two independent single-guide RNAs (sgRNAs) targeting the C-terminal PCNA interaction domain, which is essential for chromatin assembly<sup>23</sup>. Strikingly, treatment of MEFs with Cas9 and sgRNAs resulted in a similar increase in reprogramming efficiency as shRNA treatment, indicating that genetic perturbation of the endogenous *Chaf1a* locus phenocopies the effect of Chaf1a shRNAs (Fig. 3f). Sequencing of Chaf1a sgRNA-induced modifications revealed the presence of mostly biallelic (24/34) and to a lesser extent monoallelic (10/34) genome edits, suggesting that complex alterations of the *Chaf1a* coding region on one or both chromosomes can lead to hypomorphic alleles that promote iPS cell generation (Extended Data Fig. 5e). Indeed, analysis of representative iPS cell clones showed a consistent reduction but no complete loss of total Chaf1a and Chaf1b protein levels (Extended Data Fig. 3f and Supplementary Fig. 1).

Collectively, these findings demonstrate that the observed reprogramming phenotype is dependent on the duration and levels of



**Figure 3 | Enhanced reprogramming depends on optimal CAF-1 and OKSM dosage.** **a**, Comparison of reprogramming efficiency upon Chaf1a knockdown using MEFs carrying one or two copies of *Col1a1::tetOP-OKSM* and *R26-M2rtTA*. Colonies were scored at day 10 following 6 days of OKSM induction and 4 days of culture in the absence of doxycycline. **b**, Quantification of data shown in **a**. Values are the mean of biological triplicates; error bars indicate standard deviation. **c**, **d**, Effect of CAF-1 suppression on reprogramming efficiency when directly infecting MEFs with lentiviral vectors achieving medium (**c**) or high (**d**) OKSM expression levels, as determined by flow cytometry for Oct4-GFP at day 11. Values are the mean from biological triplicates; error bars indicate standard deviation. **e**, Influence of duration and degree of Chaf1a suppression on reprogramming potential of MEFs carrying doxycycline-inducible shRNA cassette (top), as determined by immunocytochemistry for Nanog at day 9. Data points represent single experiment. **f**, Comparison of reprogramming efficiencies when using shRNAs or sgRNAs targeting Chaf1a, as determined by flow cytometry for Oct4-GFP after 7 days of doxycycline exposure and 4 days of doxycycline-independent growth. Values are the mean from biological triplicates; error bars indicate standard deviation.

both CAF-1 and OKSM and that this effect can be recapitulated with multiple experimental paradigms.

### CAF-1 depletion enhances direct lineage conversion

To investigate whether CAF-1 acts as a gatekeeper of cellular identity across different cell types, we tested the effect of Chaf1a knockdown on the reprogramming potential of haematopoietic stem and progenitor cells (HSP cells) isolated from fetal livers of reprogrammable mice (Fig. 4a). Although HSP cells expressing *Renilla* shRNA gave rise to 56% and 76% Pecam<sup>+</sup> iPS-cell-like cells at days 4 and 6, respectively, suppression of Chaf1a using two independent shRNAs facilitated reprogramming towards a Pecam<sup>+</sup> state in 90% of cells at day 4 and 97% at day 6 (Extended Data Fig. 6a, b). Chaf1a knockdown cells also showed a striking elevation of Pecam expression levels at both time points, consistent with acquisition of a fully reprogrammed iPS cell state (Fig. 4b, c). Accordingly, suppression of Chaf1a gave rise to more transgene-independent colonies compared to controls (Extended Data Fig. 6c).

To assess whether CAF-1 stabilizes somatic cell identity in cell-fate conversion systems other than OKSM-mediated reprogramming, we first examined transdifferentiation of fibroblasts into induced neurons upon overexpression of the transcription factor Ascl1



in MEFs<sup>24</sup>. Transgenic MEFs harbouring doxycycline-inducible shRNAs targeting *Renilla* luciferase or *Chaf1a* in the *Col1a1* locus (see Fig. 3e) were transduced with a doxycycline-inducible *Ascl1*-expressing lentivirus and measured for induced neuron formation at day 13 (Fig. 4d). CAF-1 knockdown consistently resulted in a twofold increase ( $P = 0.0075$ ) in the number of *Map2*<sup>+</sup> neurons (Fig. 4e, f and Extended Data Fig. 6d). We next tested the effect of CAF-1 suppression during the conversion of pre-B cells into macrophages upon overexpression of the myeloid transcription factor C/EBP $\alpha$  (Fig. 4g). Consistent with a previous study<sup>25</sup>, we found that the myeloid markers *Cd14* and *Mac1* are activated in the majority of cells after 48 h of C/EBP $\alpha$  expression (Extended Data Fig. 6e–g). Remarkably, shRNA suppression of CAF-1 markedly increased *Cd14* and *Mac1* expression levels at two different time points (Fig. 4h, i) as well as the fraction of *Mac1*<sup>+</sup> and *Cd14*<sup>+</sup> cells after 24 h of C/EBP $\alpha$  induction (Extended Data Fig. 6e–g).

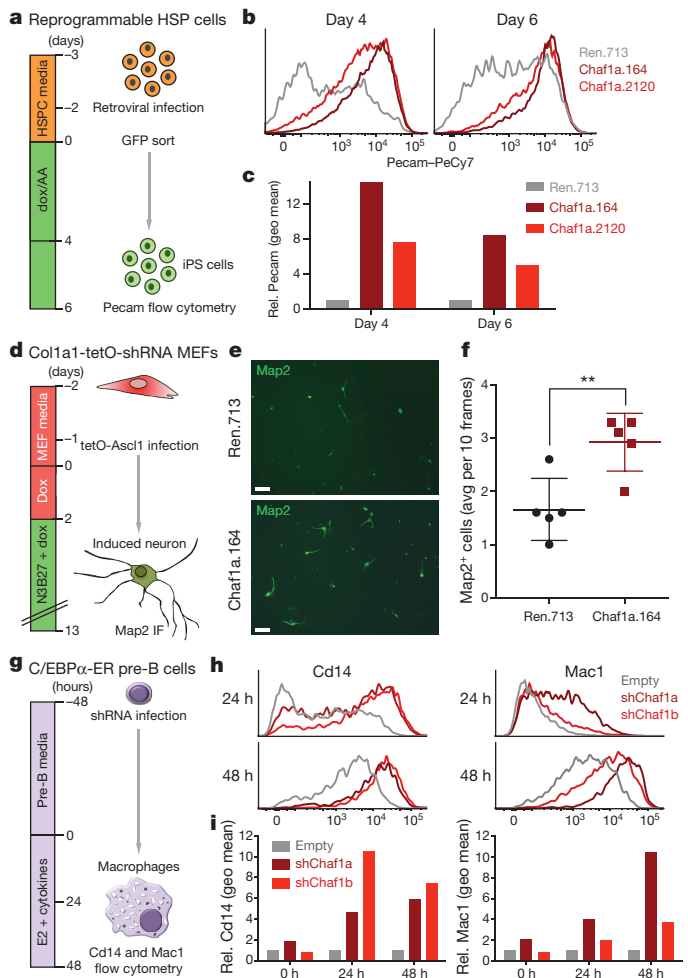
Together, these data indicate that CAF-1 suppression not only enhances the induction of pluripotency from different cell types but also facilitates cellular transdifferentiation, suggesting that CAF-1 may play a more general role in resisting transcription-factor-induced cell-fate conversions.

### CAF-1 influences chromatin accessibility

As CAF-1 functions as a histone chaperone<sup>13</sup>, we reasoned that its reduction may result in a more accessible chromatin structure and thus facilitate transcription factor binding to their target loci. To test this possibility, we first performed sonication of crosslinked chromatin sequencing (SONO-seq) analysis, which determines accessible chromatin regions based on their increased susceptibility to sonication<sup>26</sup>. We analysed bulk cultures expressing OKSM for 3 days when stable iPS cells are not yet present (Extended Data Fig. 7a), focusing on ES-cell-specific regulatory elements. Although ES-cell-specific promoter elements showed no discernible difference in accessibility ( $P = 0.51$ ), we observed a significant enrichment of SONO-seq signal at ES-cell-specific enhancer elements in CAF-1 depleted cells at day 3 of OKSM expression (Extended Fig. 7b;  $P < 2.6 \times 10^{-12}$ ).

To validate these observations with an independent, higher resolution method, we performed assay of transposase accessible chromatin-sequencing (ATAC-seq), which detects integration of the Tn5 transposase in open chromatin regions<sup>27</sup> (Extended Data Fig. 7c). Consistent with the SONO-seq data, ATAC-seq analysis of early reprogramming intermediates showed a more accessible chromatin configuration at regulatory regions including ES-cell-specific enhancers upon suppression of *Chaf1a* (Fig. 5a;  $P$  value  $< 10^{-15}$ ). Moreover, *Chaf1a* knockdown caused a significant increase in chromatin accessibility across ES-cell-specific super-enhancers at day 3 of iPS cell formation (Supplementary Table 3 and Extended Data Fig. 7d, e;  $P < 5.3 \times 10^{-16}$ ). Of note, super-enhancers linked to specialized cell types such as macrophages, lymphocytes and muscle cells were also significantly more accessible in *Chaf1a* depleted reprogramming intermediates compared to controls (Extended Data Fig. 7f). Taken together, these results suggest that CAF-1 suppression increases the permissiveness of cells to transcription-factor-induced cell-fate change by facilitating a more accessible local chromatin structure at enhancer elements.

Next, we performed ChIP-seq analysis for Sox2 at day 3 of OKSM expression in order to test our hypothesis that increased chromatin accessibility at enhancer elements influences reprogramming factor binding. Indeed, we detected an increase in Sox2 binding to ES-cell-specific regulatory elements in *Chaf1a* shRNA-treated cells compared to controls (Fig. 5b and Extended Data Fig. 8a). Approximately 90% of Sox2 binding sites were shared between *Chaf1a* knockdown and control cells, whereas 10% were unique to cells expressing either *Chaf1a* shRNA or *Renilla* shRNA (Extended Data Fig. 8b). Although *Chaf1a* knockdown cells showed slightly fewer unique Sox2 binding sites than *Renilla* knockdown cells, these sites were enriched

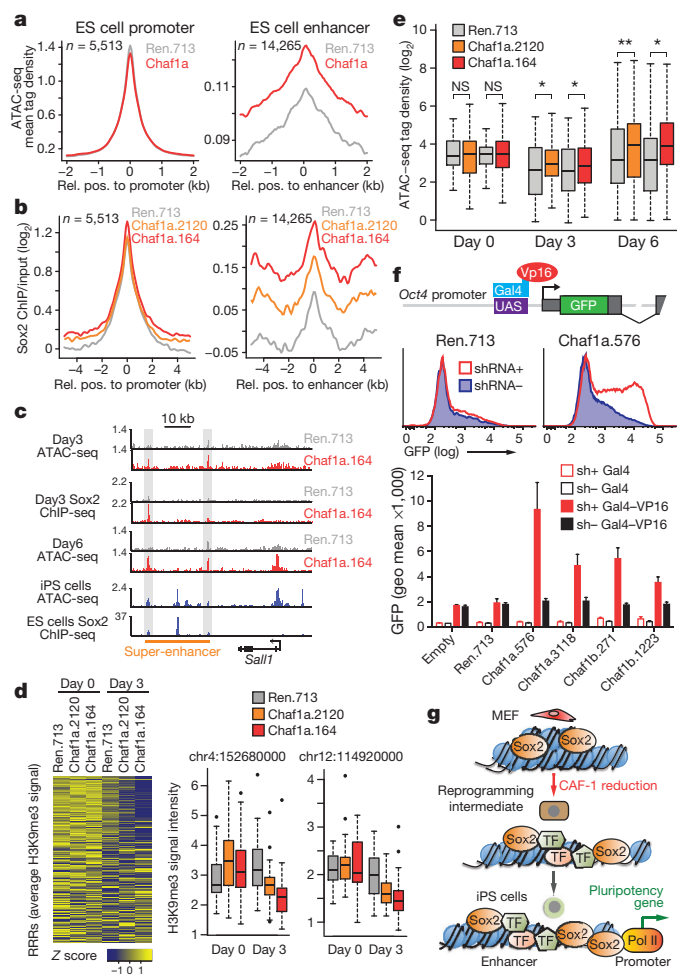


**Figure 4 | CAF-1 suppression enhances reprogramming in different cell conversion systems.**

**a**, Reprogramming of fetal haematopoietic stem and progenitor cells (HSP cells) into iPS cells. **b**, Flow cytometric analysis of *Pecam* expression during HSP cell reprogramming using indicated shRNAs. **c**, Quantification of data shown in **b**. Values represent fold-change expression differences between experimental and control samples using geometric mean. Data were obtained from one experiment using two different *Chaf1a* shRNAs. **d**, Transdifferentiation of MEFs into induced neurons. **e**, Representative image of *Map2*<sup>+</sup> induced neurons after 13 days of transdifferentiation. Scale bars, 100  $\mu$ M. **f**, Quantification of transdifferentiation efficiency ( $n = 5$  independent experiments; values are mean  $\pm$  standard deviation; unpaired  $t$ -test;  $**P = 0.0075$ ). **g**, Transdifferentiation of pre-B cells into macrophages. **h**, Activation of macrophage markers *Cd14* and *Mac1* in representative samples at indicated time points. **i**, *Cd14* and *Mac1* expression levels in indicated samples (values represent fold-change expression differences between experimental and control samples using geometric mean;  $n = 2$  independent viral transductions).

for ES-cell-specific Sox2 targets (Extended Data Fig. 8c) and ES-cell-specific super-enhancer elements (Supplementary Table 4). Of the Sox2-bound super-enhancers unique to CAF-1 knockdown cells, a subset also showed a more accessible chromatin structure by ATAC-seq analysis (for example, *Sall1*; Fig. 5c and Supplementary Table 4). Notably, Sox2 binding was also increased across lineage-specific super-enhancers when comparing *Chaf1a* knockdown cells to control at day 3 of reprogramming, consistent with the observed increase in chromatin accessibility at these elements (Extended Data Figs 7f and 8d).

Collectively, these results indicate that loss of CAF-1 contributes to reprogramming, at least in part, by increasing chromatin accessibility at pluripotency-specific enhancer elements and by promoting the binding of Sox2 to ES-cell-specific targets.



**Figure 5 | CAF-1 suppression affects chromatin dynamics and facilitates activation of pluripotency genes during iPS cell generation.**

**a**, ATAC-seq analysis of ES-cell-specific enhancers and promoters at day 3 of reprogramming. Shown are merged data for Chaf1a.164 and Chaf1a.2120 shRNA-infected cells ( $P > 0.5$  and  $P < 10^{-15}$  between Chaf1a and *Renilla* shRNAs for promoters and enhancers, respectively;  $n$  denotes number of examined promoter and enhancer elements). **b**, Sox2 ChIP-seq analysis of ES-cell-specific enhancers and promoters at day 3 of reprogramming using weak (shRNA no. 2120) and strong (shRNA no. 164) Chaf1a hairpin ( $P < 1 \times 10^{-15}$  for both shRNAs; see **a** for definition of  $n$ ). **c**, Representative ATAC-seq and Sox2 ChIP-seq peaks at the *Sal1* super-enhancer ( $y$  axis: tag density profiles). **d**, H3K9me3 ChIP-seq analysis of reprogramming-resistant regions (RRRs)<sup>29</sup> after 0 and 3 days of OKSM expression. Heatmap shows all RRRs (rows); box plots show individual RRRs between day 0 and 3 in Chaf1a knockdown cells ( $P < 0.05$  for both shRNAs). **e**, Chromatin accessibility at days 0, 3 and 6 for genes that become transcriptionally upregulated in Chaf1a shRNA-treated cells by day 6 ( $*P < 0.05$ ;  $**P < 0.01$ ). **f**, Chromatin *in vivo* assay (CiA) to directly measure effect of CAF-1 suppression on transcriptional activity of endogenous *Oct4* locus in fibroblasts upon overexpression of Gal4-VP16 fusion protein targeted to the *Oct4* promoter. **g**, Summary and model. TF, transcription factor; Pol II, RNA polymerase II.

### CAF-1 affects heterochromatin and gene expression

Considering that CAF-1 plays crucial roles not only in histone exchange but also heterochromatin maintenance<sup>19,21,28</sup>, we next examined the global distribution of the heterochromatin mark H3K9me3 during reprogramming in the presence of either Chaf1a or *Renilla* shRNAs. We did not detect significant differences in H3K9me3 levels across pluripotency-associated enhancers or transposable elements, which are typically silenced by H3K9me3 modifications (Extended Data Fig. 9a, b). Likewise, RNA-seq analysis of the same intermediates failed to show differential expression of transposable elements

(Extended Data Fig. 9c). However, we detected a local depletion of H3K9me3 at a subset of somatic heterochromatin areas termed 'reprogramming-resistant regions', which have recently been linked to the low efficiency of somatic cell nuclear transfer (SCNT)<sup>29</sup> (Fig. 5d and Extended Data Fig. 9d, e). These data suggest that CAF-1 inhibition, in concert with OKSM expression, causes local changes in this key repressive histone modification, which may prime chromatin structure for efficient transcriptional activation.

MEFs expressing OKSM and Chaf1a shRNAs upregulated a number of pluripotency-related genes (for example, *Utf1*, *Epcam*, *Nr0b1*, *Tdgf1* and *Sall4*) at day 6 of reprogramming compared to the *Renilla* shRNA control (Supplementary Table 5). Although these genes were not yet differentially expressed at day 3 of reprogramming (data not shown), chromatin associated with these genes was already more accessible, supporting the view that CAF-1 suppression may prime the genome for subsequent transcriptional activation (Fig. 5e). To exclude that transcriptional activation of these pluripotency genes is an indirect consequence of accelerated reprogramming following CAF-1 suppression, we performed a 'chromatin *in vivo* assay' (CiA)<sup>30</sup>. Specifically, we introduced CAF-1 or control shRNAs into transgenic fibroblasts carrying an array of Gal4 binding sites (UAS elements) upstream of the endogenous *Oct4* promoter and a GFP reporter in place of the *Oct4* coding region. Although expression of Gal4-VP16 alone or in combination with a control shRNA triggered weak Oct4-GFP activation, co-expression of Gal4-VP16 and independent shRNAs targeting Chaf1a or Chaf1b strongly enhanced GFP expression from the somatically silenced *Oct4* locus (Fig. 5f and Extended Data Fig. 9f). These results demonstrate that CAF-1 suppression enhances the direct transcriptional activation of the endogenous *Oct4* locus independently of OKSM-induced cell-fate changes.

### Discussion

We find that CAF-1 suppression not only enhances reprogramming towards pluripotency but also direct lineage conversion, suggesting that the study of iPS cell formation may be a valuable approach to uncover general mechanisms that safeguard somatic cell identity. Importantly, enhanced reprogramming is influenced by the degree and duration of CAF-1 suppression, consistent with the essential role of CAF-1 during cellular growth<sup>19,20,22</sup>. The identification of dose-dependent regulators of reprogramming highlights the utility of RNAi screens to achieve hypomorphic gene expression states, allowing for the detection of phenotypes that might not have been observed with complete and permanent ablation of genes. However, suboptimal reduction of CAF-1 levels may also explain our inability to detect a consistent enhancement of iPS cell formation upon CAF-1 suppression in preliminary human reprogramming experiments (data not shown). Similarly, variable CAF-1 gene dosage may account for phenotypic differences between our study and previous publications examining the function of CAF-1 in ES cells<sup>21,31</sup>. Suppression of CAF-1 in ES cells reportedly results in a decondensation of pericentric heterochromatin, the reactivation of transposable elements and cell cycle arrest and apoptosis, which differs from our observations in nascent iPS cells. An alternative explanation is that cell-intrinsic differences between early reprogramming intermediates and established pluripotent cells cause distinct phenotypes<sup>21</sup>.

It is possible that other histone chaperones compensate for reduced CAF-1 levels, thereby contributing to the observed reprogramming phenotype. In support of this notion, suppression of the CAF-1 p60 subunit in HeLa cells triggers alternative deposition of the histone variant H3.3 (ref. 32). Of interest, H3.3 deposition on chromatin has been associated with enhanced reprogramming in the context of SCNT<sup>33,34</sup>, suggesting parallels between the processes of cloning and iPS cell generation. Our finding that CAF-1 suppression leads to specific loss of H3K9me3 at reprogramming-resistant regions supports this notion. The recent insight that CAF-1 siRNA-treated ES cells revert towards a two-cell-like embryonic state more amenable to reprogramming by SCNT is consistent with this conclusion and

reinforces the view that CAF-1 may act as a general stabilizer of cell identity<sup>31</sup>. We propose a model whereby CAF-1 contributes to the maintenance of somatic cell identity by stabilizing chromatin patterns (Fig. 5g). Here, suppression of CAF-1 would trigger dilution of newly assembled nucleosomes at key enhancer elements and loosening of chromatin structure in conjunction with forced expression of cell-type-specific transcription factors. These combined changes would generate an accessible chromatin landscape for efficient transcription factor binding and activation of key target genes.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 23 February 2015; accepted 28 September 2015.**

- Lee, T. I. & Young, R. A. Transcriptional regulation and its misregulation in disease. *Cell* **152**, 1237–1251 (2013).
- Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
- Vierbuchen, T. & Wernig, M. Molecular roadblocks for cellular reprogramming. *Mol. Cell* **47**, 827–838 (2012).
- Yang, C. S., Chang, K. Y. & Rana, T. M. Genome-wide functional analysis reveals factors needed at the transition steps of induced reprogramming. *Cell Rep.* **8**, 327–337 (2014).
- Mikkelsen, T. S. *et al.* Dissecting direct reprogramming through integrative genomic analysis. *Nature* **454**, 49–55 10.1038/nature07056 (2008).
- Onder, T. T. *et al.* Chromatin-modifying enzymes as modulators of reprogramming. *Nature* **483**, 598–602 (2012).
- Rais, Y. *et al.* Deterministic direct reprogramming of somatic cells to pluripotency. *Nature* **502**, 65–70 (2013).
- Dejosez, M., Ura, H., Brandt, V. L. & Zwaka, T. P. Safeguards for cell cooperation in mouse embryogenesis shown by genome-wide cheater screen. *Science* **341**, 1511–1514 (2013).
- Qin, H. *et al.* Systematic identification of barriers to human iPSC generation. *Cell* **158**, 449–461 (2014).
- Stadtfeld, M., Maherali, N., Borkent, M. & Hochedlinger, K. A reprogrammable mouse strain from gene-targeted embryonic stem cells. *Nature Methods* **7**, 53–55 (2010).
- Zuber, J. *et al.* RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature* **478**, 524–528 (2011).
- Fellmann, C. *et al.* An optimized microRNA backbone for effective single-copy RNAi. *Cell Rep.* **5**, 1704–1713 (2013).
- Smith, S. & Stillman, B. Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication *in vitro*. *Cell* **58**, 15–25 (1989).
- Chen, J. *et al.* H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nature Genet.* **45**, 34–42 (2013).
- Soufi, A., Donahue, G. & Zaret, K. S. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* **151**, 994–1004 (2012).
- Sridharan, R. *et al.* Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1 $\gamma$  in reprogramming to pluripotency. *Nature Cell Biol.* **15**, 872–882 (2013).
- Krizhanovsky, V. & Lowe, S. W. Stem cells: the promises and perils of p53. *Nature* **460**, 1085–1086 (2009).
- Polo, J. M. *et al.* A molecular roadmap of reprogramming somatic cells into iPSCs. *Cell* **151**, 1617–1632 (2012).
- Quivy, J. P., Gerard, A., Cook, A. J., Roche, D. & Almouzni, G. The HP1-p150/CAF-1 interaction is required for pericentric heterochromatin replication and S-phase progression in mouse cells. *Nature Struct. Mol. Biol.* **15**, 972–979 (2008).
- Hoek, M. & Stillman, B. Chromatin assembly factor 1 is essential and couples chromatin assembly to DNA replication *in vivo*. *Proc. Natl Acad. Sci. USA* **100**, 12183–12188 (2003).
- Houlard, M. *et al.* CAF-1 is essential for heterochromatin organization in pluripotent embryonic cells. *PLoS Genet.* **2**, e181 (2006).
- Ye, X. *et al.* Defective S phase chromatin assembly causes DNA damage, activation of the S phase checkpoint, and S phase arrest. *Mol. Cell* **11**, 341–351 (2003).
- Roef Ben-Shahar, T. *et al.* Two fundamentally distinct PCNA interaction peptides contribute to chromatin assembly factor 1 function. *Mol. Cell. Biol.* **29**, 6353–6365 (2009).
- Chanda, S. *et al.* Generation of induced neuronal cells by the single reprogramming factor ASCL1. *Stem Cell Rep.* **3**, 282–296 (2014).
- Bussmann, L. H. *et al.* A robust and highly efficient immune cell reprogramming system. *Cell Stem Cell* **5**, 554–566 (2009).
- Auerbach, R. K. *et al.* Mapping accessible chromatin regions using Sono-seq. *Proc. Natl Acad. Sci. USA* **106**, 14926–14931 (2009).
- Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature Methods* **10**, 1213–1218 (2013).
- Huang, H. *et al.* *Drosophila* CAF-1 regulates HP1-mediated epigenetic silencing and pericentric heterochromatin stability. *J. Cell Sci.* **123**, 2853–2861 (2010).
- Matoba, S. *et al.* Embryonic development following somatic cell nuclear transfer impeded by persisting histone methylation. *Cell* **159**, 884–895 (2014).
- Hathaway, N. A. *et al.* Dynamics and memory of heterochromatin in living cells. *Cell* **149**, 1447–1460 (2012).
- Ishiyuchi, T. *et al.* Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. *Nature Struct. Mol. Biol.* (2015).
- Ray-Gallet, D. *et al.* Dynamics of histone H3 deposition *in vivo* reveal a nucleosome gap-filling mechanism for H3.3 to maintain chromatin integrity. *Mol. Cell* **44**, 928–941 (2011).
- Jullien, J. *et al.* HIRA dependent H3.3 deposition is required for transcriptional reprogramming following nuclear transfer to *Xenopus* oocytes. *Epigenetics Chromatin* **5**, 17 (2012).
- Wen, D., Banaszyński, L. A., Rosenwaks, Z., Allis, C. D. & Rafii, S. H3.3 replacement facilitates epigenetic reprogramming of donor nuclei in somatic cell nuclear transfer embryos. *Nucleus* **5**, 369–375 (2014).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank B. Kingston, C. Vakoc, M. Tolstorukov and G. Hannon for guidance and discussions, B. Bernstein, K. Plath, K. Chronis, Y. Shen and O. Tam for advice on the ATAC-seq analysis, P. Brown for providing the Dot1l inhibitor, B. Stillman for sharing the Chaf1b antibody and T. Graf for sharing the C10 cell line. We thank C. Nakada and Y. Kiyota (Nikon) for providing software to quantify iPSC cell formation and A. Huebner for help with transdifferentiation experiments. We are grateful to H. Hock and the HSCI-CRM flow cytometry core for help with flow data analysis and to W. Mallard for initial RNA-sequencing analysis. We further thank B. Ma, S. Muller, M. Weissenboeck and the IMP/IMBA Biooptics and Transgenic core facility as well as the CSF NGS laboratory for technical assistance and all members of the Hochedlinger, Zuber, Penninger, Elling, Shi and Kingston laboratories for their feedback on various aspects of this project. We thank A. Stark, A. Deaton and L. Barrett for critical reading of the manuscript. S.C. was supported by the PRCRP at the Department of Defense (CA 120212). H.Y.C. by was supported by NIH P50-HG007735. U.E. was funded by grants from IMBA and the Austrian National Foundation. S.W.L. was supported by a cancer center support grant and program project grant from the NCI and is an HHMI investigator. J.M.P. was supported by IMBA, ERC GA (number 341036) and the Innovator Award/Era of Hope Award Number W81XWH-12-1-0093. J.Z. was funded by an ERC starting grant (number 336860) and generous institutional funding from Boehringer Ingelheim. K.H. was supported by funds from the MGH, HHMI, NIH (R01 HD058013-06) and the Gerald and Darlene Jordan Chair in Regenerative Medicine.

**Author Contributions** Y. L. Jung and B. Hopfgartner contributed equally to this work. S.C., K.H., U.E. and J.Z. designed primary screens, analysed and interpreted data. S.C., J.M. and N.A. performed the arrayed screen and S.C. conducted follow-up cell biology and chromatin studies. U.E. and B.H. performed the multiplexed screen. U.E. performed validation experiments, genetic interaction assays and cell biology experiments with support from B.H., M.H. and D.W. Human reprogramming experiments were performed by S.C. and J.B.; N.T. and S.W.L. assisted in the generation of inducible *Col1a1::tetOP-Chaf1a* shRNA cell lines. S.C., A.I.B., A.B. and Y.S. performed B-cell to macrophage conversion experiments. C.E.A. and M.W. conducted MEF to induced neuron transdifferentiation experiments. Y.L.J., M.N., A.A., F.F. and P.J.P. performed bioinformatics analyses. M.H. and U.E. conducted the CiA assay with support from O.B. D.J.W. assisted with the SONO-seq experiments and H.Y.C. helped with the ATAC-seq assay. J.M., M.H. and M.Z. assisted with western blot and chromatin studies. D.T. and J.R. conducted ChIP experiments and library construction. M.S. and S.E.V. provided secondary Oct4–tdTomato MEFs. J.Z. and S.W.L. provided the arrayed library. J.Z. and P.R. designed the extended chromatin library. M.F., J.J. and B.H. generated lentiviral vectors and RNAi reagents. J.M.P. and G.A. provided intellectual support and mentoring. K.H., S.C., J.Z. and U.E. wrote the paper with input from all co-authors.

**Author Information** All SONO-seq, ATAC-seq, ChIP-seq, RNA-seq and microarray data have been deposited in the Gene Expression Omnibus database under accession number GSE66534. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to K.H. ([khochedlinger@mgm.harvard.edu](mailto:khochedlinger@mgm.harvard.edu)) or J.Z. ([johannes.zuber@imp.ac.at](mailto:johannes.zuber@imp.ac.at)).



## METHODS

**Data reporting.** No statistical methods were used to predetermine sample size. The investigators were not blinded to allocation during experiments and outcome assessment. One cell line<sup>25</sup> was used (subclone C10) in this study. Cell line identity was verified by its responsiveness to tamoxifen, resulting in conversion of cells into macrophages (quantified by surface marker expression of Mac1 and Cd14). Cells were tested for mycoplasma contamination and found to be negative.

**Plasmids.** For the pooled RNAi screen, shRNAs were expressed from the LENC vector, which has been described previously<sup>12</sup>. For the primary screen validation, timeline experiment and immunofluorescence staining, mouse shRNAs were cloned individually into LENC. For the double knockdown experiment, shRNAs were cloned into LEPC (MSCV-miR-E-PGK-Puro-IRES-mCherry). For the reprogramming dynamics experiment and chimera mouse production, shRNAs were cloned into RT3CEPIN (TRE3G-mCherry-miR-E-PGK-Puro-IRES-Neo). For reprogramming experiments with non-transgenic systems, previously published OKSM lentiviral vectors were modified to introduce promoters of different strength, which are described in the main figures.

**Cell culture and media.** Packaging cells (Platinum-E Retroviral Packaging Cell Line) for producing retroviral particles were cultured in DMEM supplemented with 15% FBS, 100 U ml<sup>-1</sup> penicillin, 100 µg ml<sup>-1</sup> streptomycin, sodium pyruvate (1 mM) and L-glutamine (4 mM) at 37°C with 5% CO<sub>2</sub>. Mouse embryonic fibroblasts (MEF) were cultured in DMEM supplemented with 15% FBS, 100 U ml<sup>-1</sup> penicillin, 100 µg ml<sup>-1</sup> streptomycin, sodium pyruvate (1 mM), L-glutamine (4 mM), L-ascorbic acid (50 µM) at 37°C with low oxygen (4.5% O<sub>2</sub>). iPS cells were derived in DMEM supplemented with 15% FBS, 100 U ml<sup>-1</sup> penicillin, 100 µg ml<sup>-1</sup> streptomycin, sodium pyruvate (1 mM), L-glutamine (4 mM), 1,000 U ml<sup>-1</sup> LIF, 0.1 mM beta-mercaptoethanol, and 50 µg ml<sup>-1</sup> ascorbic acid at 37°C with 5% CO<sub>2</sub> and 4.5% O<sub>2</sub>. For Tet-inducible OKSM expression, doxycycline was added at a concentration of 1 µg ml<sup>-1</sup> (unless indicated otherwise). iPS cells for blastocyst injection were cultured on feeders in DMEM supplemented with 13% knockout serum replacement (Gibco), 2% FBS, 100 U ml<sup>-1</sup> penicillin, 100 µg ml<sup>-1</sup> streptomycin, sodium pyruvate (1 mM), L-glutamine (4 mM), L-ascorbic acid (50 µM), 1,000 U ml<sup>-1</sup> LIF, beta-mercaptoethanol (50 µM), MEK inhibitor (PD0325901, 1 µM) and GSK3 inhibitor (CHIR99021, 3 µM) at 37°C with 5% CO<sub>2</sub>. Conventional reprogramming media consisted of DMEM supplemented with 15% FBS, 100 U ml<sup>-1</sup> penicillin, 100 µg ml<sup>-1</sup> streptomycin, sodium pyruvate (1 mM), L-glutamine (4 mM), 1,000 U ml<sup>-1</sup> LIF, 0.1 mM beta-mercaptoethanol unless otherwise noted. For some experiments, media was supplemented with MEK inhibitor (1 µM), GSK3 inhibitor (3 µM), Dot1l inhibitor (1 µM) or ascorbate (50 µg ml<sup>-1</sup>). Reprogrammable MEFs containing either one or two copies of the *Col1a1::tetOP-OKSM*, *Oct4-GFP* and *Rosa26 M2rtTA* alleles<sup>10</sup> were derived from E13.5 embryos. MEFs were prepared after carefully excluding internal organs, heads, limbs and tails. Tissues were chopped into small clumps using scalpels and trypsin and subsequently expanded in MEF medium at low O<sub>2</sub> (4%). MEFs were frozen at passage 0 upon derivation and used at passages 1–3 for all downstream transduction and reprogramming experiments. MEFs were generally cultured at low O<sub>2</sub> (4%) and supplemented with ascorbate to prevent replicative senescence before OKSM induction. Reprogramming experiments were initiated at low oxygen levels during doxycycline induction and completed at normal oxygen levels (20%) for experiments using miR-E vectors. MiR-30 assays were performed under normal oxygen levels. To generate *Col1a1::tetOP-miR30-tRFP-Ren.713* and *Col1a1::tetOP-miR30-tRFP-Chaf1a.164* shRNA knock-in MEFs, miR30-based shRNAs targeting Chaf1a.164 or Ren.713 were cloned into a targeting vector as previously described<sup>35</sup> except that the GFP reporter was replaced with a turbo RFP reporter. ES cells harbouring the *R26-M2rtTA* allele were targeted with these constructs and mice were generated by blastocyst injection. MEFs were harvested using standard protocols.

HSP cells were isolated from fetal livers of the same mid-gestation reprogrammable transgenic embryos used for MEFs derivation, dissociated by vigorous pipetting with a 1 ml tip, filtered using a 35 µm nylon mesh, followed by red blood cell lysis, and cultured in RPMI/FBS media supplemented with stem cell factor (SCF), IL3 and IL6 and transduced as indicated in the schematic (Fig. 4a).

**Arrayed shRNA library preparation and screening.** Single shRNA clones were picked from the master library at CSHL, arrayed in 12 × 96-well plates and sequence-verified individually using miR-30 backbone primers. An additional 200 unmatched clones were re-picked and sequenced to allow maximum coverage of the library. Reprogrammable MEFs carrying the OKSM inducible cassette and constitutive rtTA (*Col1a1::tetOP-OKSM*; *R26-M2rtTA*) were seeded at 10<sup>4</sup> cells per well in 96-well plates in duplicates and infected with the corresponding retroviral virus particles freshly produced and filtered. 48 h post-transduction, MEFs from each row of the 96-well plate were trypsinized and transferred to 6-well dishes coated with 0.2% gelatin in standard reprogramming media supplemented with doxycycline (2 µg ml<sup>-1</sup>) and G418 at 0.2 mg ml<sup>-1</sup> for the first 6

days of OKSM expression. Doxycycline was withdrawn at day 12, allowing stable iPS cells to form. iPS cell colonies were then stained for alkaline phosphatase expression using the Vector Red Alkaline Phosphatase Substrate Kit (VectorLabs) according to the supplier's protocol, and plates were scanned using a Perfection V500 Photo scanner (Epson). To determine relative reprogramming efficiencies (Fig. 1c and Supplementary Table 1), automated counting of iPS cell colonies was performed using the image-processing software CL-Quant (Nikon) and a custom algorithm provided by NIKON. Data were normalized to Ren.713 control.

**Pooled shRNA library preparation and screening.** A miR-E-based chromatin library comprising 5,049 sequence-verified shRNAs targeting 615 known and predicted chromatin regulators was constructed by subcloning pools of sequence-verified miR-30 shRNAs into pLENC and combining them at equimolar concentrations into one pool<sup>12</sup>. This pool was transduced into MEFs carrying the *Col1a1::tetOP-OKSM* and *R26-M2rtTA* alleles, as well as a Pou5f1-EGFP reporter (termed Oct4-GFP) under conditions predominantly yielding a single retroviral integration in the genome. To generate a large number of independent biological replicates, primary MEFs from 4 triple transgenic embryos were transduced with the entire pool of 5,049 shRNAs in 12 independent replicates at a representation of >100 cells per shRNA, yielding a total of 48 replicates (see Fig. 1b). After 36 h, MEFs were treated with 0.5 mg ml<sup>-1</sup> G418 for 3 days and 0.25 mg ml<sup>-1</sup> G418 for an additional 3 days. MEFs from each replicate were plated at densities of 500,000 cells per 15 cm dish 3 or 6 days post-transduction, and induced with doxycycline (1 µg ml<sup>-1</sup>) for 7 days in medium containing serum and LIF, supplemented with ascorbate (50 µg ml<sup>-1</sup>). After passaging for an additional 4 days in doxycycline-free ES cell media, Oct4-GFP-expressing cells were sorted from each replicate using a FACSAriaIII (BD Bioscience).

Genomic DNA from infected MEFs (3d after infection) and sorted Oct4-GFP iPS cells from each replicate was isolated using proteinase K lysis, followed by two rounds of phenol extraction using PhaseLock tubes (5prime) and isopropanol precipitation. Templates for deep-sequencing were generated by PCR amplification of shRNA guide strands using primers that tag the product with standard Illumina adapters (p7+loop, CAAGCAGAAGACGGCATACGA[4-nt barcode]TAGTGAAGCCACAGATGT; p5+PGK, AATGATACGCGACACCACCGATGGATGTGGAATGTGTGC GAGG). For each sample, DNA was amplified in 12 parallel 50 µl PCR reactions using Encyclo Polymerase (Evrogen). PCR products were combined for each sample, precipitated and purified on a 2% agarose gel. Samples were analysed on an Illumina High Seq 2500 and sequenced using a primer that reads in reverse into the guide strand (mirEEcoR1Seqprimer, TAGCCCCCTTGAAGTCCGAGGCAGTAGGCA). Sequence processing was performed using a customized Galaxy platform. In all 96 iPS cell samples (48 biological replicates, 3 or 6 days knockdown before OKSM expression) the normalized reads of each shRNA were divided by the normalized reads in MEFs 3 days after viral transduction, and the resulting ratio was used to calculate a score for each shRNA in each replicate (default score = 0; score = 1 if ratio >1, score = 3 if ratio >10). Scores of each shRNA in 48 replicates were added separately for the day 3 and day 6 time point, yielding a sum score to estimate the overall enrichment of each shRNA over all replicates. All shRNA sequences and primary results from the arrayed and the multiplexed screen are provided in Supplementary Tables 1 and 2, respectively. shRNAs are identified by numbers (e.g. Ren.713, Chaf1a.164), defined as the 5' nucleotide of the guide binding site in the target transcript at the time of shRNA design.

**Retrovirus production, transduction of MEFs and derivation of iPS cells.** Retroviral constructs were introduced into Platinum-E Retroviral Packaging cells using calcium phosphate transfection or lipofection as previously described<sup>36</sup>. shRNAs were transduced into primary MEFs carrying single copies of the *Col1a1::tetOP-OKSM* and *R26-M2rtTA* alleles as well as the Oct4-EGFP reporter. For some experiments, Oct4-tomato knock-in MEFs were used; the *Oct4-tomato* allele was generated equivalently to the *Oct4-GFP* allele<sup>37</sup>. For transduction, 180,000 cells were plated per well in a 6-well dish; all vectors were transduced in biological triplicate. After 36 h, transduced cells were selected with 0.5 mg ml<sup>-1</sup> G418 for 3 days and 0.25 mg ml<sup>-1</sup> G418 for an additional 3 days. Then 3 days after shRNA transduction, infected cells were washed with PBS (1 ×) and trypsinized with Trypsin-EDTA (1 ×) and 20,000 cells were plated into a 6-well. OKSM expression was induced for 7 days and cells were cultured in DMEM supplemented with 15% FBS, 100 U ml<sup>-1</sup> penicillin, 100 µg ml<sup>-1</sup> streptomycin, sodium pyruvate (1 mM), L-glutamine (4 mM), 1,000 U ml<sup>-1</sup> LIF, 0.1 mM beta-mercaptoethanol, 50 µg ml<sup>-1</sup> sodium ascorbate and 1 µg ml<sup>-1</sup> doxycycline at 37°C with low oxygen (4.5% O<sub>2</sub>) and 5% CO<sub>2</sub>. After 7 days of OKSM expression, cells were cultured for an additional 4 days without doxycycline to withdraw OKSM transgene expression at 37°C with 5% CO<sub>2</sub>, ambient oxygen. Following trypsinization, cells were analysed for Oct4-GFP expression using a FACS BD LSRFortessa (BD Biosciences), data were analysed using FlowJo.

**Phenotypic characterization of iPS cells.** Alkaline phosphatase activity was measured using an enzymatic assay for alkaline phosphatase (VECTOR red alkaline phosphatase (AP) substrate kit) according to the manufacturer's protocol. Nanog immunohistochemistry of iPS cell colonies was performed as previously described<sup>15</sup> using anti-Nanog antibody (ab80892, Abcam) at a dilution of 1:500. Cells were permeabilized with 0.2% Triton-X before blocking and antibody incubation.

**Chimera production and germline transmission assays.** Triple transgenic MEFs were reprogrammed as described, using a tetOP-inducible shRNAmiR expression vector RT3CEPIN (TRE3G-mCherry-mirE-PGK-Puro-IRES-Neo). Oct4-GFP<sup>+</sup> iPS cells generated with experimental shRNAs were sorted on day 7 of OKSM transgene expression. iPS cells were plated on feeders and cultured in DMEM supplemented with 13% knockout serum replacement, 2% FBS, 100 U ml<sup>-1</sup> penicillin, 100 µg ml<sup>-1</sup> streptomycin, sodium pyruvate (1 mM), L-glutamine (4 mM), L-ascorbic acid (50 µM), 1,000 U ml<sup>-1</sup> LIF, beta-mercaptoethanol, MEK inhibitor (PD0325901, 1 µM) and GSK3 inhibitor (CHIR99021, 3 µM) at 37 °C with 5% CO<sub>2</sub>. Polyclonally derived iPS cells were microinjected into B6 albino blastocysts to allow identification of chimaeras based on coat colour markers. Male chimaeras were mated to B6 albino females to allow identification of germline transmission based on coat colour.

**Double knockdown assay.** Triple transgenic reprogrammable MEFs were transduced with shRNA expressed from LENC as previously described and cultured in MEF media. 3 days after retroviral infection, cells were sorted for mCherry expression and 40,000 cells were re-plated per well of a 6-well dish. On the next day, cells were infected with the corresponding second shRNA expressed from LENC. 24 h later, cells were cultured in DMEM supplemented with 15% FBS, 100 U ml<sup>-1</sup> penicillin, 100 µg ml<sup>-1</sup> streptomycin, sodium pyruvate (1 mM), L-glutamine (4 mM), 1,000 U ml<sup>-1</sup> LIF, 0.1 mM beta-mercaptoethanol and 1 µg ml<sup>-1</sup>, 50 µg ml<sup>-1</sup> sodium ascorbate, and 1 µg ml<sup>-1</sup> doxycycline at 37 °C with lox oxygen (4.5% O<sub>2</sub>). 36 h after the second shRNA transduction, cell culture media was supplemented with 0.5 mg ml<sup>-1</sup> G418 for 3 days and 0.25 mg ml<sup>-1</sup> G418 for an additional 3 days to ensure double infection. After 7 days of OKSM transgene induction, cells were cultured in ES cell medium for an additional 4 days without doxycycline to select for transgene-independent colonies at 37 °C with 5% CO<sub>2</sub>. Cells were analysed for Oct4-GFP expression using a FACS BD LSRFortessa (BD Biosciences). The effect of double knockdown of targets on iPS cell formation was determined by calculating the ratio of Oct4-GFP<sup>+</sup> to Oct4-GFP<sup>-</sup> cells at day 11 relative to an empty vector control.

**Analysis of reprogramming dynamics.** Triple transgenic reprogrammable MEFs carrying the *Col1a1::tetOP-OKSM* and *R26-M2rtTA* alleles as well as a Oct4-GFP reporter were reprogrammed in replicate wells as previously described. Starting on day 4 of OKSM transgene expression, cells were analysed for Oct-GFP expression in 24 h intervals using a BD LSRFortessa (BD Biosciences). In addition, 20% of cells were replated and cultured under doxycycline-free ES cell conditions. After 13 days, cells were analysed using a FACS BD LSRFortessa to determine the minimum time required for the establishment of transgene-independent iPS cells.

To determine Nanog expression dynamics, triple transgenic reprogrammable MEFs were reprogrammed in independent wells and analysed every 24 h. Starting on day 4 of OKSM transgene expression, cells were trypsinized with trypsin-EDTA (1×), washed with PBS (1×) and fixed with paraformaldehyde (PFA) (4%) for 30 min. Afterwards, cells were washed with PBS (1×) and stored at 4 °C. After 11 days of reprogramming, cells were stained with anti-Nanog antibody (rabbit polyclonal, 1:400, Abcam) and analysed using a FACS BD LSRFortessa (BD Biosciences).

Pecan staining of reprogramming intermediates was performed as previously described<sup>18</sup>. All samples were analysed on a MACSQuant fluorescence cytometer (Miltenyi).

**CRISPR/Cas9 editing of MEFs.** sgRNAs targeting the *Chaf1a* locus were cloned into a lentiviral vector harbouring the wild type Cas9 coding region, an sgRNA expression cassette, and a Thy 1.1 reporter transgene. Successfully transduced cells were purified by FACS using Thy1.1 expression, cultured for 7 days to allow for genome editing to occur and induced with doxycycline for one week before measuring the fraction of Oct4-GFP<sup>+</sup> cells at day 11. Single Oct4-GFP<sup>+</sup> iPS cells were then plated to generate clonal iPS cell cultures for PCR amplification of CRISPR/Cas9-induced genomic modifications, followed by Sanger sequencing. sgRNAs were PCNA-1: GAAGCGCATTAAGGCAGAAA and PCNA-2: TTGGAGCCTGCGGAGTCTT.

**Transdifferentiation assays.** Induced neurons were generated as described in the experimental scheme (Fig. 4d). CAF-1 or *Renilla* RNAi inducible transgenic MEFs were transduced with Ascl1-inducible lentivirus, exposed to doxycycline 24 h post-induction, cultured in MEF media for the first 48 h and switched to serum-free neuronal media (N3B27) supplemented with doxycycline for an

additional 11 days. Cultures were fixed and stained for Map2 as previously described<sup>24</sup>.

Pre-B cells (C10 line)<sup>25</sup> were cultured in RPMI Medium, 10% charcoal stripped FBS (Invitrogen), 2 mM L-glutamine, 100 U ml<sup>-1</sup> penicillin, 100 µg ml<sup>-1</sup> streptomycin, 55 µM beta-mercaptoethanol. Pre-B cells were transduced with lentiviral pLKO vectors obtained from the Broad Institute's RNAi consortium (empty vector 'null control' or vector carrying stem-loop shRNAs targeting Chaf1a and Chaf1b subunits). Following selection of transduced cells with puromycin, cells were seeded at 10<sup>6</sup> cells per ml and supplemented with oestradiol (E2) and macrophage cytokines (IL3 and CSF) to induce macrophage transdifferentiation as previously described<sup>25</sup>. All time points were analysed for Cd14 and Mac1 expression by flow cytometry on the same day.

**Quantitative RT-PCR.** RNA was extracted (Qiagen RNeasy mini kit) and reverse transcribed (GE illustra ready-to-go RT-PCR beads) according to the supplier's instruction. Quantitative PCR was performed using SybrGreen and a BIO-RAD CFX connect cyclor. Primers used were: b-Act-F: GCTGTATCCCCTCCATCGTG; b-Act-R: CACGGTTGGCCTTAGGGTTCAG; Ube2i-R: GGCAAACCTCTTCGCTTGCTCGGAC; Ube2i-F: ATCCTTCTGGCACAGTGTGCCTGTCC; Chaf1b-R: GGCTCCTTGCTGTCATTCATCTTCCAC; Chaf1b-F: CACCGCCGTGAGGATCTGGAAGTTGG; Chaf1a-R: GTGTCTTCTCAACTTTCTCTCTTG; Chaf1a-F: CGCGGACAGCCGCGCCGTGGATTGC.

**SDS-PAGE and western blot analysis.** Whole-cell lysates from reprogramming intermediates were run on 4–20% gradient SDS-polyacrylamide gels and transferred to nitrocellulose membrane (Bio-Rad) by standard methods. Membranes were blocked for 1 h in 5% non-fat dry milk in 1×TBS with 0.05% Tween-20 (TBST), rinsed, and incubated with primary antibody diluted in 3% BSA in TBST overnight at 4 °C. The following primary antibodies were used: anti-Chaf1a (sc-10206, Santa Cruz), anti-Chaf1b (sc-393662, Santa Cruz), anti-TBP (ab818, Abcam), anti-Ube2i (4786, Cell signaling), anti-PCNA (D3H8P, Cell signaling), HRP conjugate anti-actin (AC-15, Sigma). Blots were washed in TBST, incubated with HRP-conjugated secondary antibodies for semi-quantitative western blot analysis and IRdye 800CW or IRdye 680RD for quantitative westerns, as indicated. Secondary antibodies for both methods were incubated in 5% milk in TBST for 1 h at room temperature (except for anti-β-actin-peroxidase antibody, which was incubated for 15 min), and washed again. HRP signal was detected by Enhanced Chemiluminescence (Perkin Elmer). Fluorescent infrared signal was detected using LI-COR Odyssey imaging system.

**ATAC-seq chromatin assay.** To generate ATAC-seq libraries, 50,000 cells were used and libraries were constructed as previously described<sup>27</sup>. Briefly, cells were washed in PBS twice, counted and nuclei were isolated from 100,000 cells using 100 µl hypotonic buffer (10 mM Tris pH 7.4, 10 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.1% NP40) to generate two independent transposition reactions. Nuclei were split in half and treated with 2.5 µl Tn5 Transposase (Illumina) for 30 min at 37 °C. DNA from transposed nuclei was then isolated and PCR-amplified using barcoded Nextera primers (Illumina). Library quality control was carried out using high-sensitivity DNA bioanalyzer assay and qubit measurement and sequenced using paired-end sequencing (PE50) on the Illumina Hi-Seq 2500 platform.

**SONO-seq and ChIP-seq chromatin assays.** For all ChIP experiments, 107 reprogramming intermediates were collected per library. Chromatin precipitation assays were performed as previously described<sup>38</sup> using goat polyclonal anti-Sox2 antibody (AF2018, R&D). Briefly, cells were cross-linked on plate in 1% methanol-free formaldehyde and snap-frozen in liquid nitrogen until processed. Nuclei were isolated using 1 ml of cell lysis buffer (20 mM Tris pH 8, 85 mM KCl, 0.5% NP40 and 1×HALT protease inhibitor cocktail), resuspended in nuclear lysis buffer (10 mM Tris-HCl pH 7.5, 1% NP40, 0.5% Na deoxycholate, 0.1% SDS, 1×HALT protease inhibitor cocktail) and sonicated using optimized pulses of a Branson sonifier (1 min ON/OFF pulses for 5 cycles) for ChIP-seq libraries and S220 Covaris sonicator (settings: duty cycle 5%, intensity 6, cycles/burst 200, pulse length 60 s, 20 cycles, 8 °C) for SONO-seq input preparations. Sonifications were verified for both methods using the 2100 Bioanalyzer. Immunoprecipitations were carried out by first adjusting salt concentration in sheared chromatin to 167 mM NaCl and adding antibodies (6 µg of Sox2 antibody) and incubated for 3–4 h at 4 °C. 50 µl Protein G dynabeads (Invitrogen) were prepared for each IP reaction by washing 2 to 3 times in ChIP dilution buffer (16.7 mM Tris-HCl pH 8.1, 167 mM NaCl, 0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA) and added for one additional hour to pull-down bound chromatin. Bead complexes were washed 6 times in RIPA buffer (20 mM Tris-HCl pH 8.1, 1 mM EDTA, 140 mM NaCl, 1% Triton X-100, 0.1% SDS, 0.1% Na deoxycholate), then twice with RIPA buffer with high salt concentration (500 mM), then twice in LiCl buffer (10 mM Tris-HCl pH 8.1, 1 mM EDTA, 1% DOC, 1% NP40, 250 mM LiCl) and twice in TE buffer. Complexes were then eluted and reverse crosslinked in 50 µl ChIP



elution buffer (10 mM Tris-HCl pH 8, 5 mM EDTA, 300 mM NaCl, 0.1% SDS) and 8  $\mu$ l of reverse crosslinking buffer (250 mM Tris-HCl pH 6.5, 1.25 M NaCl, 62.5 mM EDTA, 5 mg ml<sup>-1</sup> proteinase K, 62.5  $\mu$ g ml<sup>-1</sup> RNase A) by incubation at 65 °C for 6 h. DNA was isolated using Ampure SPRI beads and yield quantified using Qubit fluorometer.

ChIP-seq libraries were constructed from 10 ng of immunoprecipitated DNA using the NEBNext ChIP-seq library prep reagent set for Illumina (New England Biolabs), following the supplier's protocol. Briefly, purified DNA was end-repaired and dA-tailed. Following subsequent ligation of sequencing adaptors, ligated DNA was size-selected to isolate fragments in the range of 300–550 bp in length using Egel. Adaptor-ligated fragments were enriched in a 14-cycle PCR using Illumina multiplexing primers. Libraries were purified, analysed for correct size distribution using dsDNA High Sensitivity Chips on a 2100 Bioanalyzer (Agilent), pooled and submitted for single-end 50 bp Illumina GAII high-throughput sequencing.

**SONO-seq bioinformatics analysis.** The reads were aligned to the mouse genome (mm 9) using Bowtie with the unique mapping option<sup>39</sup>. The smoothed tag density profiles were generated using *get.smoothed.tag.density* function of the SPP R package<sup>40</sup> with a 100-bp Gaussian kernel, a 50-bp step and library size normalization. The positions of promoters and enhancers in ES cells were obtained from a publicly available data set<sup>41</sup>. To assess the significance of the difference in the enrichment values between CAF-1 and *Renilla* knockdown samples, a paired Wilcoxon rank sum test was used.

**ATAC-seq bioinformatics analysis.** The reads were aligned to mm9 using BWA version 0.7.8 with -q 5 -l 32 -k 2 and paired option<sup>42</sup>. Non-primary mapping, failed QC, duplicates and non-paired reads were filtered. If one paired-end was mapped to one chromosome and the other end was mapped to a different chromosome, the read was not included. The reads aligned to chrM were also removed. Only uniquely mapped reads were used. The read density profiles were generated using 150-bp windows with a 20-bp step and were normalized by the library size. For the comparison between Chaf1a shRNA and *Renilla* shRNA samples, the read density profiles were further normalized using the mean values of all annotated promoters from mm 9. For meta-analysis, the reads from Chaf1a.164 and Chaf1a.2120 knockdown samples were merged. The coordinates of promoters and enhancers in ESCs and MEFs were obtained from a publicly available data set<sup>41</sup>. The coordinates of the super-enhancers for the meta-gene plot were used from a recently published data set<sup>43</sup>. Each super-enhancer region (with 5-kb margins) was divided into 101 bins and the tag density signals were averaged in each bin. Significantly enriched regions were detected using Hotspot<sup>44</sup> with FDR = 0.01. A one-sided paired Wilcoxon rank sum test was used for the comparison in the enrichment values between CAF-1 and *Renilla* knockdown samples. To classify the genomic locations of the peaks (promoters, coding exons, introns, intergenic regions, 5' UTR and 3' UTR), the annotations for mm 9 were downloaded from UCSC (<https://genome.ucsc.edu/cgi-bin/hgTables>). The differential sites between CAF-1 and *Renilla* knockdown samples were identified using DiffBind with  $P = 0.05$  for the consensus ATAC-seq peaks after normalization with TMM (trimmed mean normalization method)<sup>45</sup>. DiffBind uses statistical routines developed in edgeR<sup>46</sup>. A one-sided paired Wilcoxon rank sum test was used for the comparison in the enrichment values between CAF-1 and *Renilla* knockdown samples.

**Sox2 ChIP-seq bioinformatics analysis.** The reads were aligned and tag densities profiles were generated as in SONO-seq analyses. The log<sub>2</sub>-fold enrichment profiles were generated using *get.smoothed.enrichment.mle* in the SPP R package. The profiles were normalized by the background-scaling method using non-enriched regions. A paired Wilcoxon rank sum test was used for the comparison in the Sox2 enrichment values between Chaf1a and *Renilla* knockdown samples. For Sox2 peak comparison between CAF-1 and *Renilla* knockdown samples, reads were first subsampled to make the sequencing depth the same for each condition (number of peaks called tends to increase for greater sequencing depth). The significantly enriched peaks compared to input were detected using the SPP *find.binding.positions* function with default parameters. The overlapped peaks were compared with a margin of 200 bp. For unique peaks, we first identified the peaks that were present only in one condition (CAF-1 or *Renilla* knockdown) and compared the enrichment values (input-subtracted tag counts) between CAF-1 and *Renilla* knockdown. If the ratios between the enrichment values were greater than two-fold, we considered the peaks as 'unique' for one of the conditions. We used Sox2 ChIP-seq data in ES cells from publicly available data sets<sup>47</sup> and analysed data in the same as described above.

**H3K9me3 ChIP-seq bioinformatics analyses.** ChIP-seq data were mapped to the mouse genome (mm 9) with Bowtie 0.12.7 (ref. 39) allowing up to 3 mismatches, retaining uniquely mapping reads. To assess H3K9me3 signal distribution genome-wide, we divided the genome in 5-kb intervals, and for each interval, we calculated the ratio of RPM normalized signal in the IP and input

samples. Intervals with less than 10 reads in the input samples (~10% of all) were excluded from further analyses due to low coverage. Intervals overlapping specific regions were extracted using the bedtools suite<sup>48</sup>. RRR annotations were obtained from ref. 29, and signal across all included 5-kb intervals was averaged. For H3K9me3 enrichment over transposable element (TE) bodies, we used the mm 10 genome version, as this release contains the most recent TE annotations. We extracted the genomic regions corresponding to TE families annotated in the mm10 RepeatMasker track in the UCSC genome browser (<http://genome.ucsc.edu/>), and calculated the normalized read counts in IP to input samples for each family. Due to the repetitive nature of TEs, we further validated all results considering reads that map to multiple (up to 10,000) positions in the genome, and scaling read counts by the number of valid alignments. This threshold for multiple mapping positions was chosen as it was previously shown to approximate results obtained allowing unlimited mapping positions, but at a significantly improved computation speed<sup>49</sup>. In all analyses, signal estimates based on uniquely mapping reads and based on reads mapping to multiple genomic positions produced similar results.

**Gene expression analysis using microarrays.** The microarray data were preprocessed using Affymetrix Expression Console version 1.3.0.187 and normalized by the RMA procedure. The limma Bioconductor package was used to select differentially expressed genes with false discover rate (FDR)  $\leq 0.05$  and at least two-fold change<sup>50</sup>. We performed functional analysis with gene set enrichment analysis (GSEA)<sup>51</sup> using the limma moderated  $t$ -statistic to rank the genes.

**Association of ATAC-seq changes with transcriptional changes.** ATAC-seq peaks were separately called for CAF-1 and *Renilla* knockdown at days 0, 3 and 6 as described above. To determine which genes from Supplementary Table 5 may be affected by altered ATAC-seq signals, we incorporated long-range interaction data between promoters and enhancers based on ChIA-PET analysis in ES cells<sup>52</sup>. If there was no matched pair from the ChIA-PET tables, the regions proximal to the TSSs of genes (<4kb) were taken. The regions were overlapped with the union ATAC-seq peaks of each conditions. For the overlapped peaks, the enrichment values (log<sub>2</sub> tag counts) were compared between CAF-1 and *Renilla* knockdown samples with two-sided paired Wilcoxon rank sum test.

**RNA-seq analysis of genes and transposable element bioinformatics analysis.** RNA sequencing data was first pre-processed using Reapeer<sup>53</sup> to remove any Illumina adaptor sequences and computationally depleted of ribosomal RNA sequences (GenBank identifiers: 18S, NR\_003278.3; 28S, NR\_003279.1; 5S, D14832.1; and 5.8S, K01367.1) using Bowtie 0.12.7 allowing three mismatches<sup>39</sup>. For protein-coding gene expression analyses, pre-processed data was mapped to the mouse genome (mm 10) using Bowtie 0.12.7 (ref. 39) allowing three mismatches, and retaining uniquely mapping reads. Mouse transcript annotations were obtained from RefSeq, and reads corresponding to the exonic regions of each gene were calculated using a custom python script. For overlapping genes, reads corresponding to overlapping regions were divided equally. Gene differential expression was analysed using the DESeq R package<sup>54</sup>.

For TE expression analyses, data was mapped to the mm 10 genome with 0 mismatches and considering reads that map to up to 10,000 genomic positions as in ChIP sequencing analyses. We then calculated the number of reads corresponding to TE regions annotated by the UCSC RepeatMasker track, scaling by the number of valid alignments for each read. Scaled reads for each TE family were summed, and normalized as RPM. Heatmaps were generated using the gplots R package, and differential expression analyses were performed using the DESeq R package<sup>54</sup>. Comparisons of RNA-sequencing results from analyses based on uniquely mapping reads, and based on reads mapping to multiple genomic positions, showed very similar results.

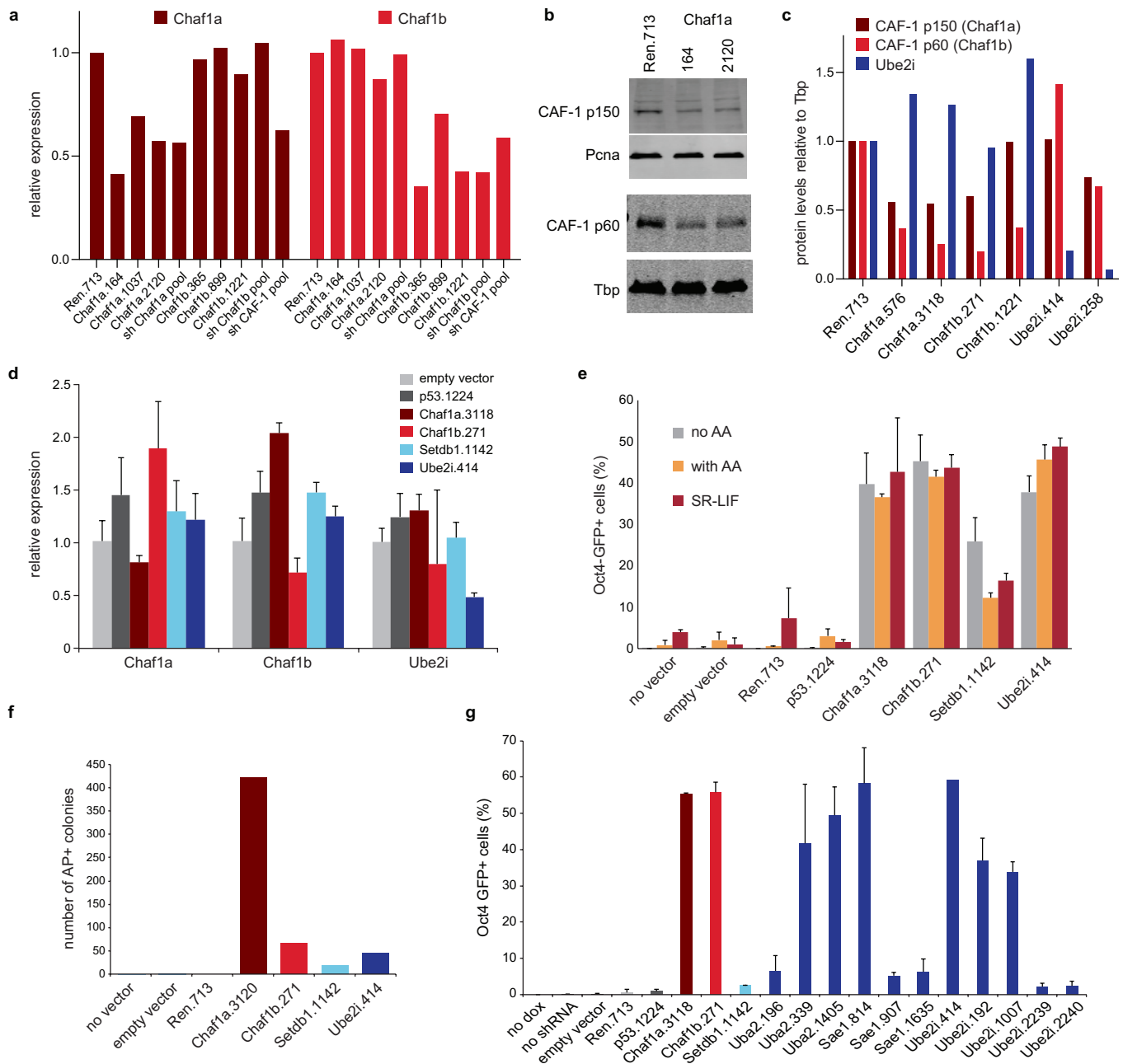
**Statistical analyses.** Unpaired Student  $t$ -test was used for statistical analysis in replicates of cell biology experiments. All error bars represent s.d. of independent biological replicates as indicated. A  $P$  value of  $<0.05$  was considered statistically significant. Numbers of replicate experiments ( $n$ ) are shown in figure legends. All graphs with no error bars represent  $n = 1$ . To assess significant differences in signal enrichment at ESC promoters, enhancers or super-enhancers by SONO-seq, ATAC-seq and ChIP-seq analysis upon CAF-1 knockdown or *Renilla* knockdown, a paired Wilcoxon rank sum test was used, where it is assumed that populations do not follow normal distributions. To identify differential ATAC-seq peaks between CAF-1 and *Renilla* knockdown samples, negative binomial models were used.

**Chromatin in-vivo assay.** CiA transgenic MEFs carrying an array of Gal4 binding sites (UAS elements) upstream of the endogenous *Oct4* promoter and a GFP reporter in place of the *Oct4* coding region<sup>30</sup> SV40-large T antigen and sub-cloned. Two clonal derivatives of these MEFs were infected with retroviral LENC vectors expressing Chaf1a, Chaf1b or *Renilla* shRNA (Fig. 5 and data not shown). Cells were subsequently transduced with lentiviral vectors expressing either



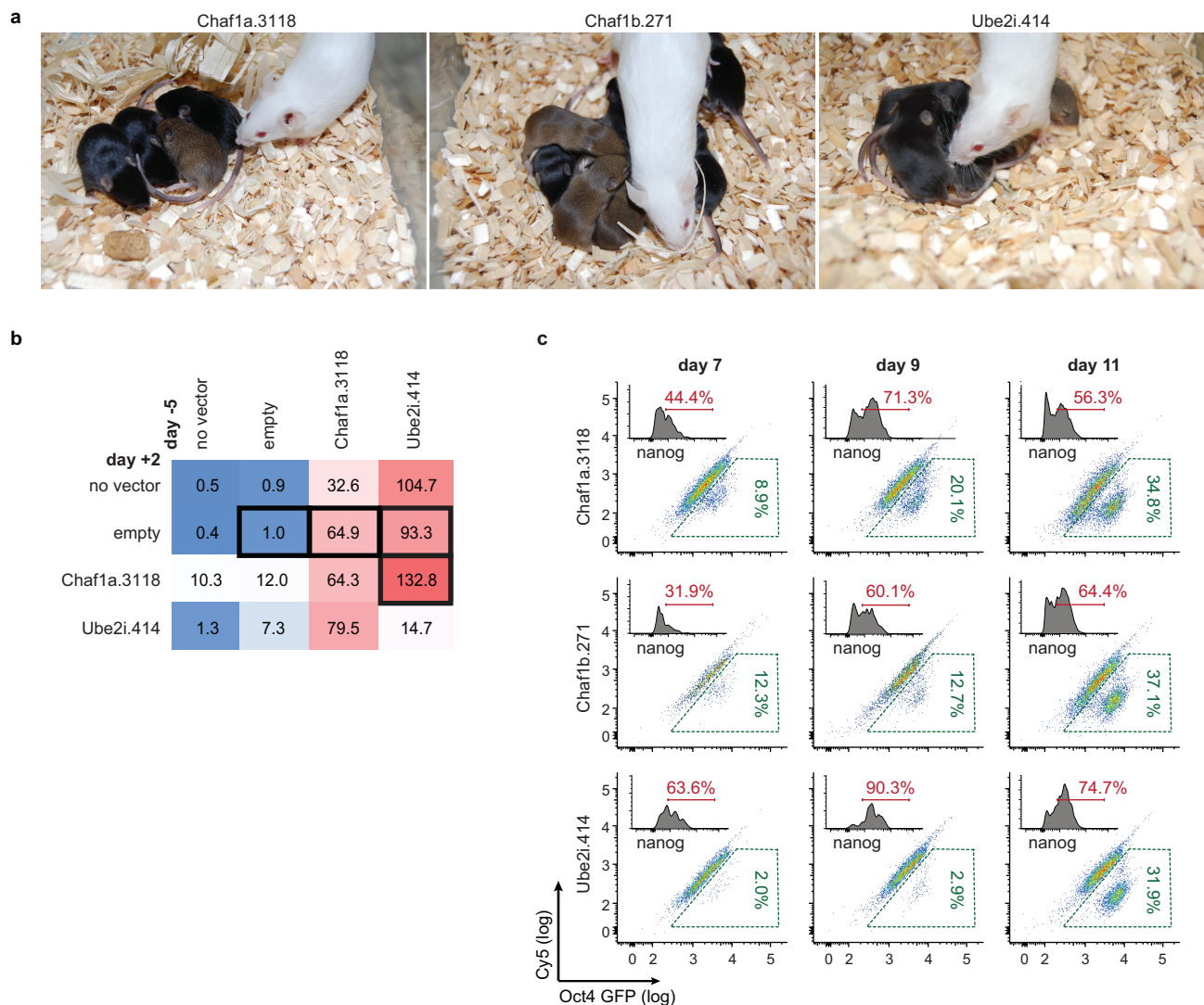
Gal4 alone or Gal4–VP16 in combination with a puromycin resistance cassette. Following drug selection, Oct4–GFP expression was measured by flow cytometry after 10 days.

35. Premysirut, P. K. *et al.* A rapid and scalable system for studying gene function in mice using conditional RNA interference. *Cell* **145**, 145–158 (2011).
36. Zuber, J. *et al.* Toolkit for evaluating genes required for proliferation and survival using tetracycline-regulated RNAi. *Nature Biotechnol.* **29**, 79–83 (2011).
37. Lengner, C. J. *et al.* Oct4 expression is not required for mouse somatic stem cell self-renewal. *Cell Stem Cell* **1**, 403–415 (2007).
38. Bernstein, B. E. *et al.* Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**, 169–181 (2005).
39. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
40. Kharchenko, P. V., Tolstorukov, M. Y. & Park, P. J. Design and analysis of ChIP-seq experiments for DNA-binding proteins. *Nature Biotechnol.* **26**, 1351–1359 (2008).
41. Shen, Y. *et al.* A map of the *cis*-regulatory sequences in the mouse genome. *Nature* **488**, 116–120 (2012).
42. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
43. Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
44. Sabo, P. J. *et al.* Discovery of functional noncoding elements by digital analysis of chromatin structure. *Proc. Natl Acad. Sci. USA* **101**, 16837–16842 (2004).
45. Ross-Innes, C. S. *et al.* Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389–393 10.1038/nature10730 (2012).
46. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
47. Marson, A. *et al.* Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* **134**, 521–533 (2008).
48. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
49. Pezic, D., Manakov, S. A., Sachidanandam, R. & Aravin, A. A. piRNA pathway targets active LINE1 elements to establish the repressive H3K9me3 mark in germ cells. *Genes Dev.* **28**, 1410–1428 (2014).
50. Smyth, G. K. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**, 1–25 (2004).
51. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
52. Downen, J. M. *et al.* Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* **159**, 374–387 (2014).
53. Davis, M. P., van Dongen, S., Abreu-Goodger, C., Bartonicek, N. & Enright, A. J. Kraken: a set of tools for quality control and analysis of high-throughput sequence data. *Methods* **63**, 41–49 (2013).
54. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).



**Extended Data Figure 1 | Validation of hits from chromatin-focused shRNA screens.** **a**, Quantitative RT-PCR analysis to confirm suppression of Chaf1a and Chaf1b expression with miR-30-based vectors from arrayed screen. Sh Chaf1a pool, sh Chaf1b pool and sh CAF-1 pool denote pools of shRNAs targeting either Chaf1a, Chaf1b or both. **b**, Western blot analysis to confirm knockdown of CAF-1 components using the top-scoring miR-30-based shRNAs from arrayed screen (see Supplementary Fig. 1 for full scans). **c**, Quantification of data shown in Fig. 1f. **d**, Quantitative RT-PCR analysis confirming knockdown with top-scoring miR-E-based shRNAs targeting Chaf1a, Chaf1b or Ube2i from the multiplexed screen. Error bars show s.d. from biological triplicates. RNA and protein were extracted from reprogrammable MEFs 72 h after doxycycline induction in panels **a**–**d**. **e**, Suppression of CAF-1 components, Ube2i

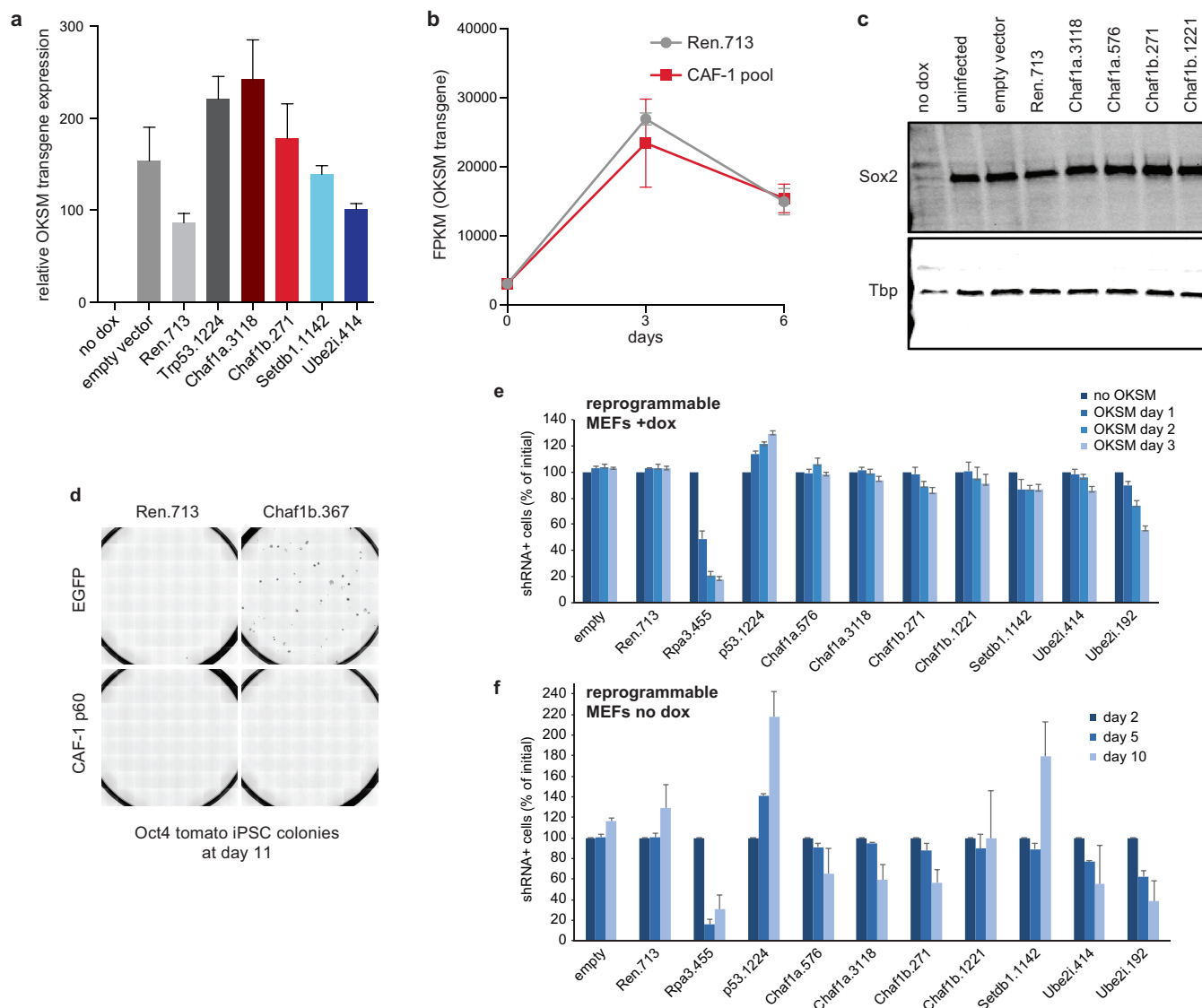
and Setdb2 enhances reprogramming in the presence or absence of ascorbic acid (AA) as well as in serum replacement media containing LIF (SR-LIF). Oct4-GFP+ cells were scored by flow cytometry on day 11 after 7 days of OKSM induction and 4 days of transgene-independent growth. Error bars show s.d. from biological triplicates. **f**, Number of doxycycline-independent, alkaline phosphatase (AP)-positive colonies emerging two weeks after plating 10,000 reprogrammable MEFs carrying shRNA vectors against indicated targets and cultured in serum replacement media containing 2i (SR-2i),  $n = 1$  experiment. **g**, Effect of suppressing SUMO E2 ligase Ube2i, E1 ligases Sae1 and Uba2 on iPS cell formation. Shown is fraction of Oct4-GFP+ cells at day 11 (7 days of OKSM induction, 4 days of transgene-independent growth). Error bars depict s.d. from biological triplicates.



**Extended Data Figure 2 | Germline transmission of iPS cells, genetic interaction of shRNA hits and effect of CAF-1 or Ube2i suppression on reprogramming dynamics.** **a**, Germline transmission of agouti chimaeras generated from iPS cells using doxycycline-inducible shRNA vectors targeting Chaf1a, Chaf1b or Ube2i. Germline transmission was determined by scoring for agouti coat colour offspring upon breeding chimaeras with albino females. Germline transmission was observed in 8/8, 4/4 and 6/8 cases for Chaf1a iPS-cell-derived chimaeras, in 7/7, 4/4, 7/7 and 9/9 cases for Chaf1b iPS-cell-derived chimaeras, and in 5/5, 7/7

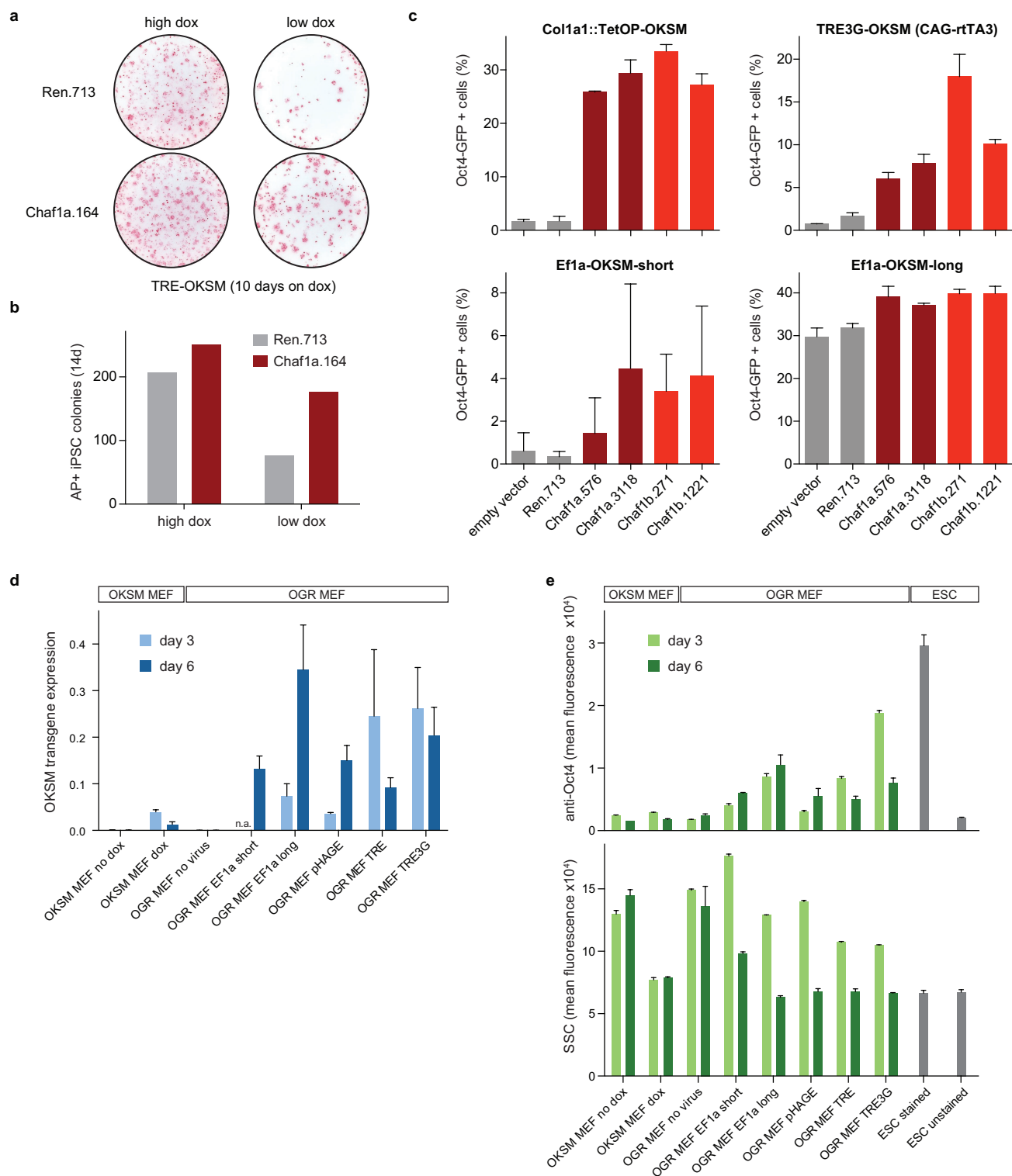
and 5/5 cases for Ube2i iPS-cell-derived chimaeras. **b**, Table summarizing effects of co-suppressing pairs of targets on emergence of Oct4-GFP<sup>+</sup> cells, shown as the ratio of Oct4-GFP<sup>+</sup> to Oct4-GFP<sup>-</sup> cells relative to an empty vector control. Experiment equivalent to Fig. 2b except that second shRNAs were transduced two days after induction of reprogramming. **c**, Representative FACS plots showing effects of Chaf1a/b or Ube2i suppression on emergence of Oct4-GFP<sup>+</sup> cells at days 7, 9 and 11 of OKSM expression. Histogram plots show fraction of Nanog<sup>+</sup> cells within Oct4-GFP<sup>+</sup> cells.





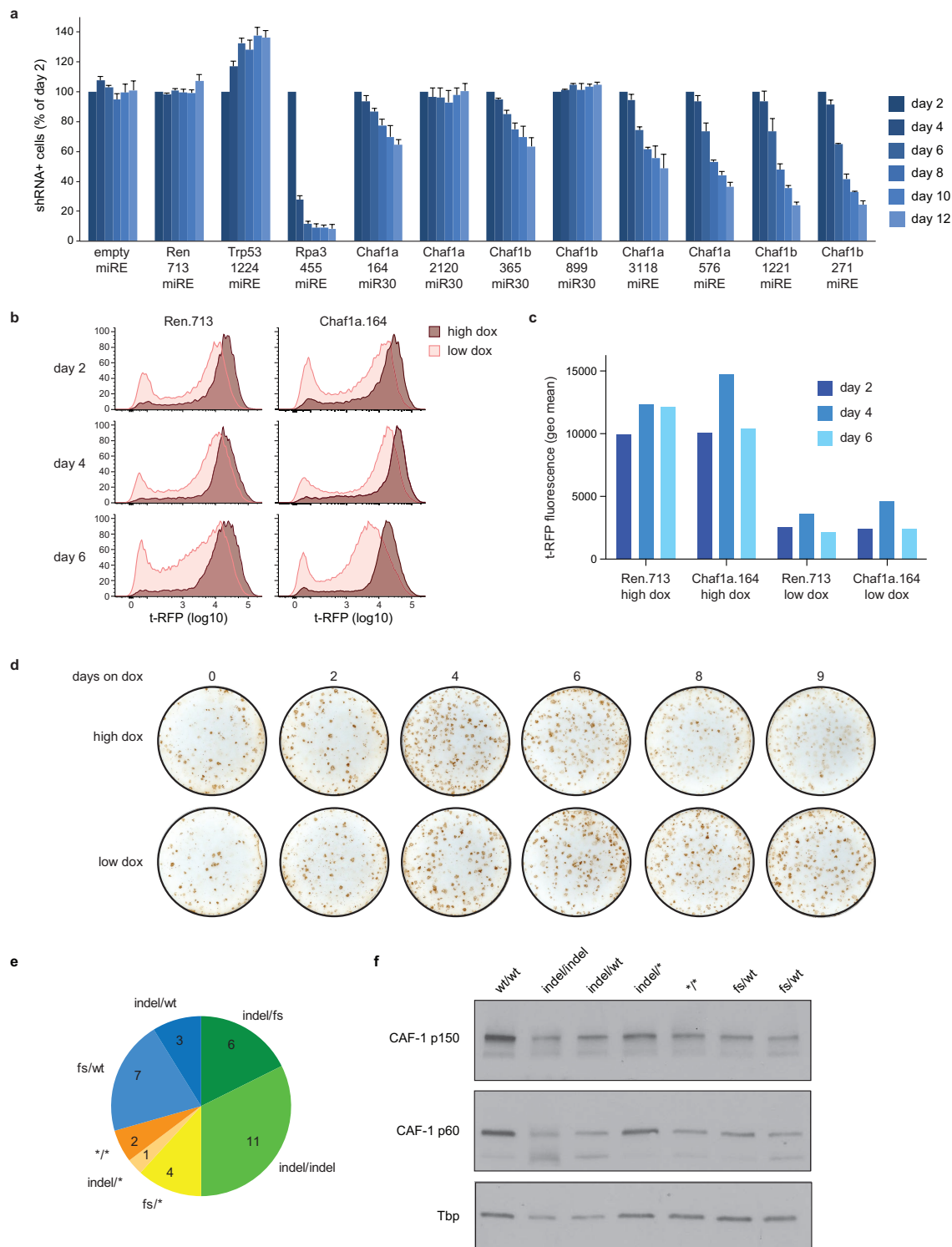
**Extended Data Figure 3 | Effect of CAF-1 suppression on OKSM levels and cellular growth, and shRNA rescue experiment.** **a**, Quantitative RT-PCR for transgenic OKSM expression using reprogrammable MEFs transduced with indicated shRNA vectors. Error bars show s.d. from biological triplicates. **b**, RNA-seq analysis of OKSM transgene expression in reprogrammable MEFs transduced with *Renilla* and Chaf1a shRNAs and exposed to doxycycline for 0, 3 or 6 days. Error bars indicate s.d. from biological triplicates. **c**, Western blot analysis for Sox2 and Tbp (loading control) in reprogrammable MEFs transduced with shRNA vectors targeting *Renilla* (Ren.713) or different CAF-1 components and exposed to doxycycline for 3 days (see Supplementary Fig. 1 for full scans). The same membrane was probed with anti-CAF-1 p150 and anti-CAF-1 p60 antibody to confirm knockdown (data not shown). **d**, Rescue experiment to demonstrate specificity of Chaf1b.367 shRNA vector. Reprogrammable

MEFs carrying Oct4–tomato knock-in reporter were infected with lentiviral vectors expressing either EGFP or human CAF-1 p60 (CHAF1B) before transducing cells with *Renilla* or Chaf1b.367 shRNAs and applying doxycycline for 6 days. Colonies were counted at day 11. Note that CAF-1 p60 overexpression attenuates enhanced reprogramming elicited by Chaf1b suppression. **e**, **f**, Competitive proliferation assay between shRNA vector-infected and non-infected reprogrammable cells using indicated shRNAs in the presence or absence of doxycycline (OKSM expression). Note that CAF-1 suppression does not substantially affect the proliferation potential of reprogrammable MEFs after 1–3 days of doxycycline (OKSM) induction while it impairs the long-term growth potential of uninduced MEFs. Data were normalized to cell counts in ‘no OKSM’ condition for **e** and ‘day 2’ time point for **f**. Error bars show s.d. from biological triplicates.



**Extended Data Figure 4 | Confirmation of CAF-1 reprogramming phenotype with alternative transgenic and non-transgenic vector systems.** **a**, Alkaline phosphatase (AP)-positive, transgene-independent iPSC cell colonies at day 14 following transduction of R26-M2rtTA MEFs with tetO-STEMCCA lentiviral OKSM expression vector and either Chaf1a.164 or Ren.713 shRNA vectors and treatment with high ( $2 \mu\text{g ml}^{-1}$ ) or low ( $0.2 \mu\text{g ml}^{-1}$ ) doses of doxycycline for 10 days. **b**, Quantification of data shown in **a**. Experiment was performed at 3 different plating densities ( $n=1$  experiment per density), representative data are shown. **c**, Comparison of reprogramming efficiencies between Col1a1::tetOP-OKSM; R26-M2rtTA reprogrammable MEFs and wild-type MEFs infected directly with OKSM-expressing lentiviral vectors containing either a strong Ef1a full-length

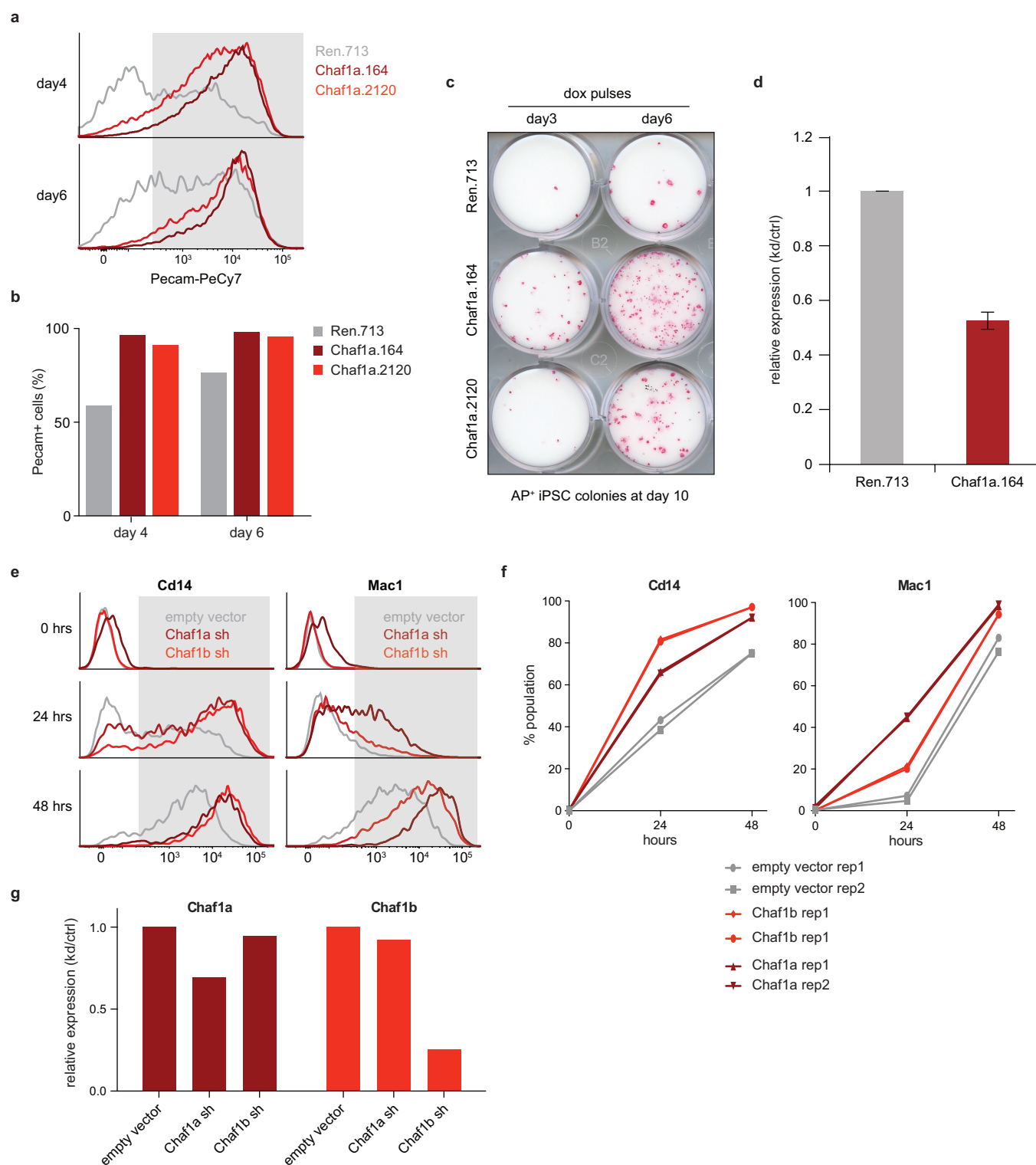
promoter (Ef1a-OKSM long) or a weaker truncated promoter (Ef1a-OKSM short). TRE3G-OKSM is a lentiviral vector with a strong promoter, whose activity is downregulated over time upon infection of CAGS-rtTA3 transgenic MEFs (see below). Error bars show s.d. from biological triplicates. **d**, Quantitative RT-PCR data showing variability in OKSM expression levels over time using different vector systems. Cells were analysed after 3 and 6 days of infection (lentiviral vectors) or doxycycline exposure (reprogrammable MEFs). Error bars show s.d. from biological triplicates. OGR MEF, transgenic MEFs carrying Oct4-GFP and CAGS-rtTA3 alleles. **e**, Quantification of Oct4 protein levels by intracellular flow cytometry (top) and cellular granularity/complexity by side scatter (SSC) analysis of indicated samples (bottom). Error bars show s.d. from biological triplicates.



**Extended Data Figure 5 | Effects of CAF-1 dose on NIH3T3 growth and reprogramming potential.** **a**, Competitive proliferation assay to determine effect of indicated Chaf1a and Chaf1b shRNA vectors on long-term growth potential of immortalized NIH3T3 cell line. Cells were infected with indicated constructs and the fraction of shRNA vector-positive cells was measured by flow cytometry at different time points. Data were normalized to cell counts at day 2 post-transduction. Rpa3.455, validated control shRNA targeting the broadly essential replication protein A3. Error bars show s.d. from biological triplicates. **b**, Histogram plots of MEFs harbouring R26-M2rtTA allele and either *Col1a1::tetOP-miR30-tRFP-Ren.713* or *Col1a1::tetOP-miR30-tRFP-Chaf1a.164* shRNA knock-in allele after transduction with pHAGE (Efla-OKSM) lentiviral vector and exposure of cells to different doses of doxycycline for 2, 4 and 6 days. Low doses of doxycycline ( $0.2 \mu\text{g ml}^{-1}$ ) result in lower expression of the shRNA

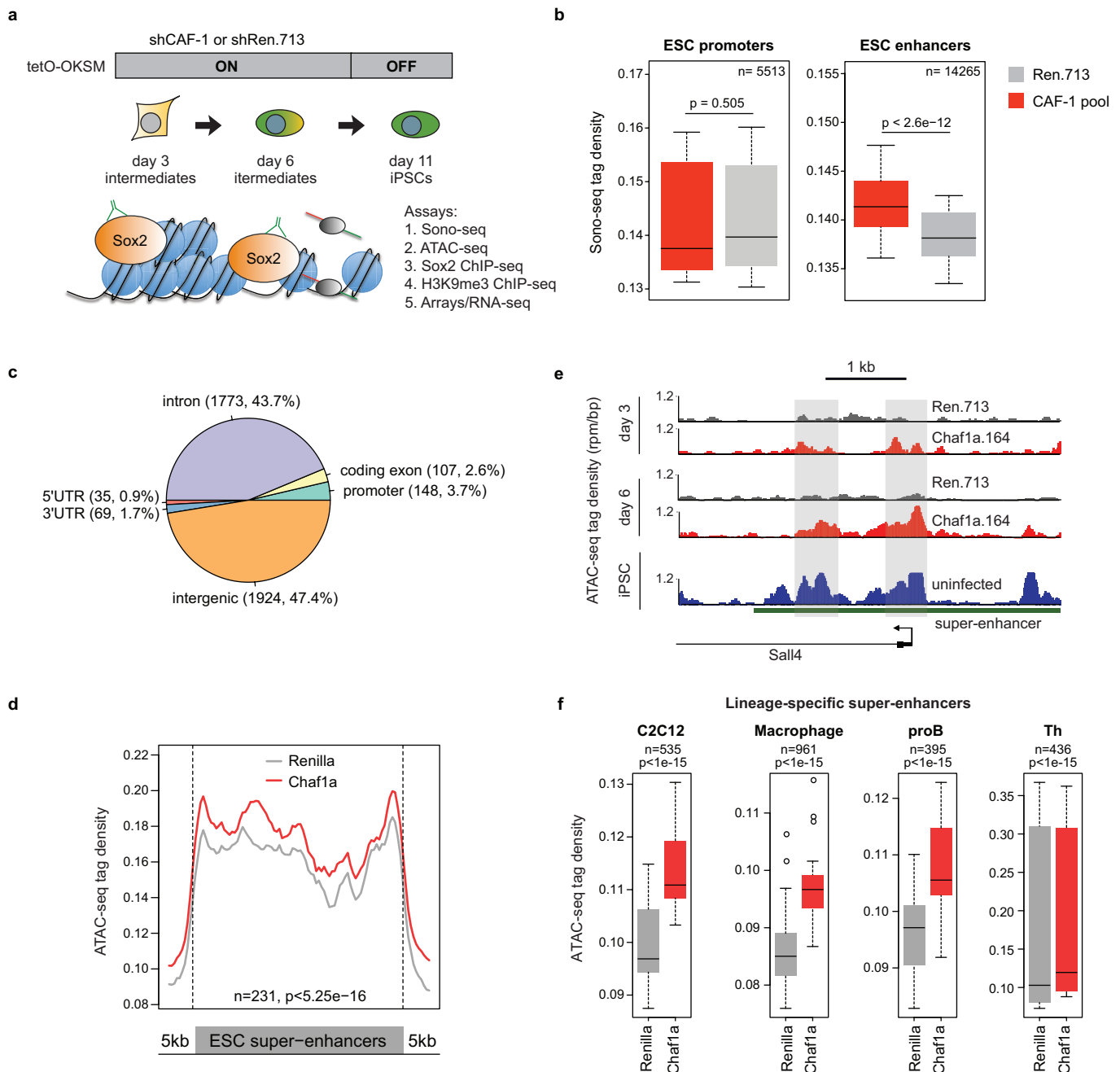
miR cassettes than high doses of doxycycline ( $2 \mu\text{g ml}^{-1}$ ). **c**, Quantification of data shown in **b** using the geometric mean ( $n = 1$  experiment for 3 indicated time points). **d**, Reprogramming efficiency of *Col1a1::tetOP-miR30-tRFP-Chaf1a.164*; R26-M2rtTA MEFs infected with pHAGE (Efla-OKSM) vector and induced with high ( $2 \mu\text{g ml}^{-1}$ ) or low ( $0.2 \mu\text{g ml}^{-1}$ ) doses of doxycycline for indicated number of days before scoring for Nanog<sup>+</sup> iPS cells by immunocytochemistry on day 9. **e**, Classification of CRISPR/Cas9-induced mutations by sequence analysis of representative iPS cell clones (wt, wild type; indel, insertion/deletion; fs, frameshift; \*, point mutation). **f**, Western blot analysis for CAF-1 subunits p150 and p60 in 6 representative iPS cell clones after CRISPR/Cas9-induced modifications of the *Chaf1a* locus (see Supplementary Fig. 1 for full scans). Wt/wt samples show unmodified wild-type control samples.





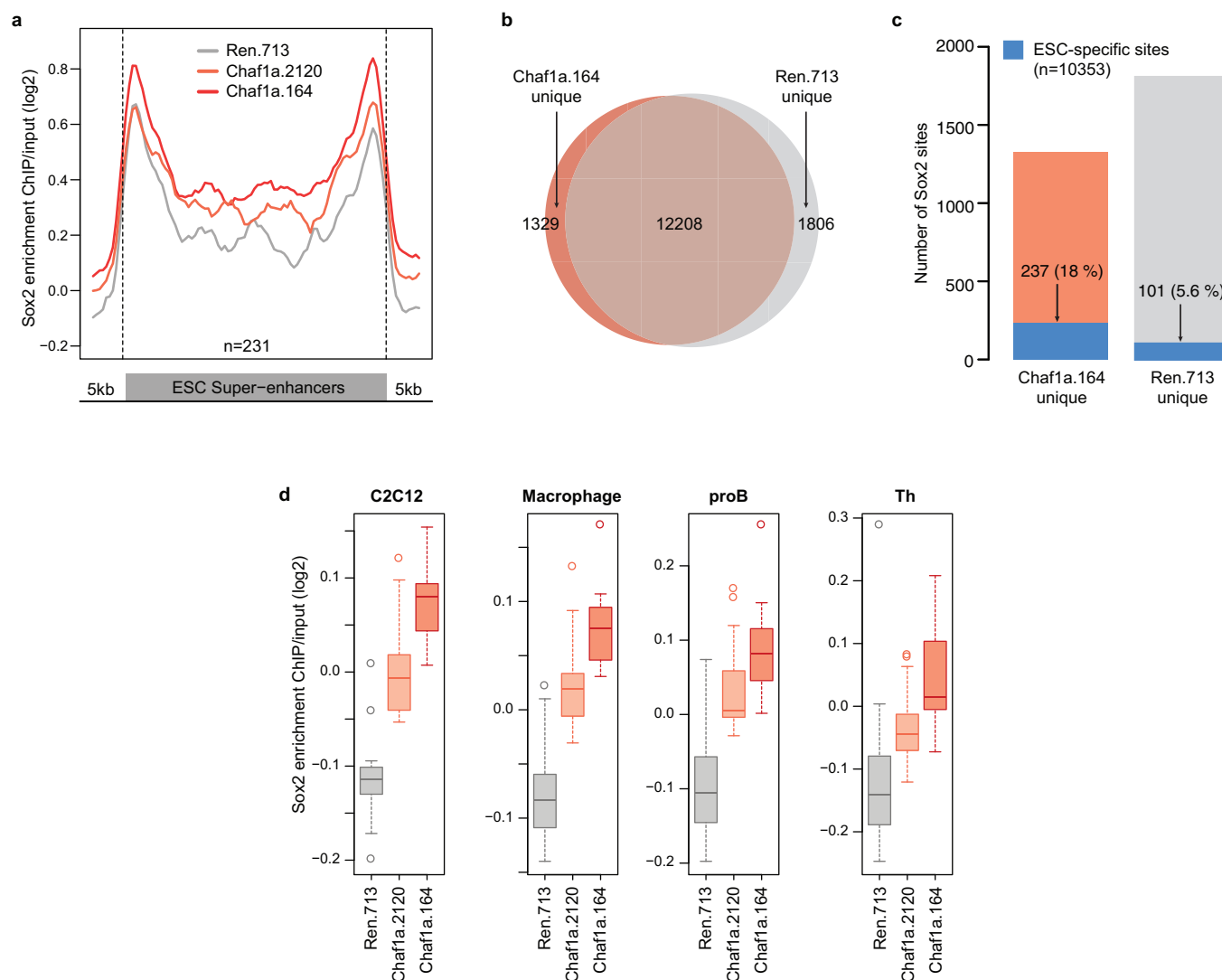
**Extended Data Figure 6 | Effect of CAF-1 suppression on HSP cell reprogramming and transdifferentiation.** **a**, Gating strategy for determining Pecam<sup>+</sup> fraction (shaded area) in panel **b**; data identical to Fig. 4c. **b**, Quantification of the fraction of Pecam<sup>+</sup> cells at day 4 and day 6 of reprogramming. Data obtained from one experiment using two different Chaf1 shRNAs. **c**, Transgene dependence assay during the reprogramming of haematopoietic stem and progenitor cells (HSP cells) into iPS cells in the presence of Chaf1a or *Renilla* shRNAs. Doxycycline pulses were given for 3 or 6 days and alkaline phosphatase (AP)-positive colonies were scored at day 10. **d**, Quantitative RT-PCR analysis of Chaf1a expression to confirm knockdown after 3 days of doxycycline induction,

that is, coexpression of shRNAmiR and Ascl1 ( $n = 4$  independent infections of the same *Col1a1::tetOP-Chaf1a.164* shRNA MEF line; mean value  $\pm$  s.d.). **e**, Gating strategy for determining Cd14<sup>+</sup> and Mac1<sup>+</sup> fractions (shaded area) shown in **f**; data identical to Fig. 4g. Positive gates were based on untreated (0 h) control cells. **f**, Quantification of the fraction of Cd14<sup>+</sup> and Mac1<sup>+</sup> cells at 0, 24 and 48 h of transdifferentiation using indicated CAF-1 shRNA or empty control vector ( $n = 2$  independent infections; rep, replicate). **g**, Quantitative RT-PCR analysis of Chaf1a and Chaf1b expression to confirm knockdown in transduced pre-B cell line before induction of transdifferentiation (kd/ctrl, knockdown/empty vector control;  $n = 1$  experiment, representative of 2 independent infections).



**Extended Data Figure 7 | CAF-1 suppression promotes chromatin accessibility at enhancer elements.** **a**, Experimental outline and assays (SONO-seq, ATAC-seq, Sox2 ChIP-seq, H3K9me3 ChIP-seq, microarrays and RNA-seq) to dissect effect of CAF-1 suppression on chromatin accessibility, transcription factor binding, heterochromatin patterns and gene expression. Assays were performed either in early reprogramming intermediates (day 3) or throughout the reprogramming time course (ATAC-seq and gene expression). **b**, SONO-seq analysis of CAF-1 knockdown and control cells at day 3 of reprogramming to determine accessible chromatin regions across promoters ( $n = 5,513$ ) and ES-cell-specific enhancers ( $n = 14,265$ ). CAF-1 shRNA vectors Chaf1a.164, Chaf1a.2120, Chaf1b.365 and Chaf1b.1221 were pooled for this experiment. **c**, ATAC-seq peak distribution across different genomic features. Shown is classification of peaks that are gained in CAF-1 knockdown cells compared to *Renilla*. **d**, ATAC-seq analysis of

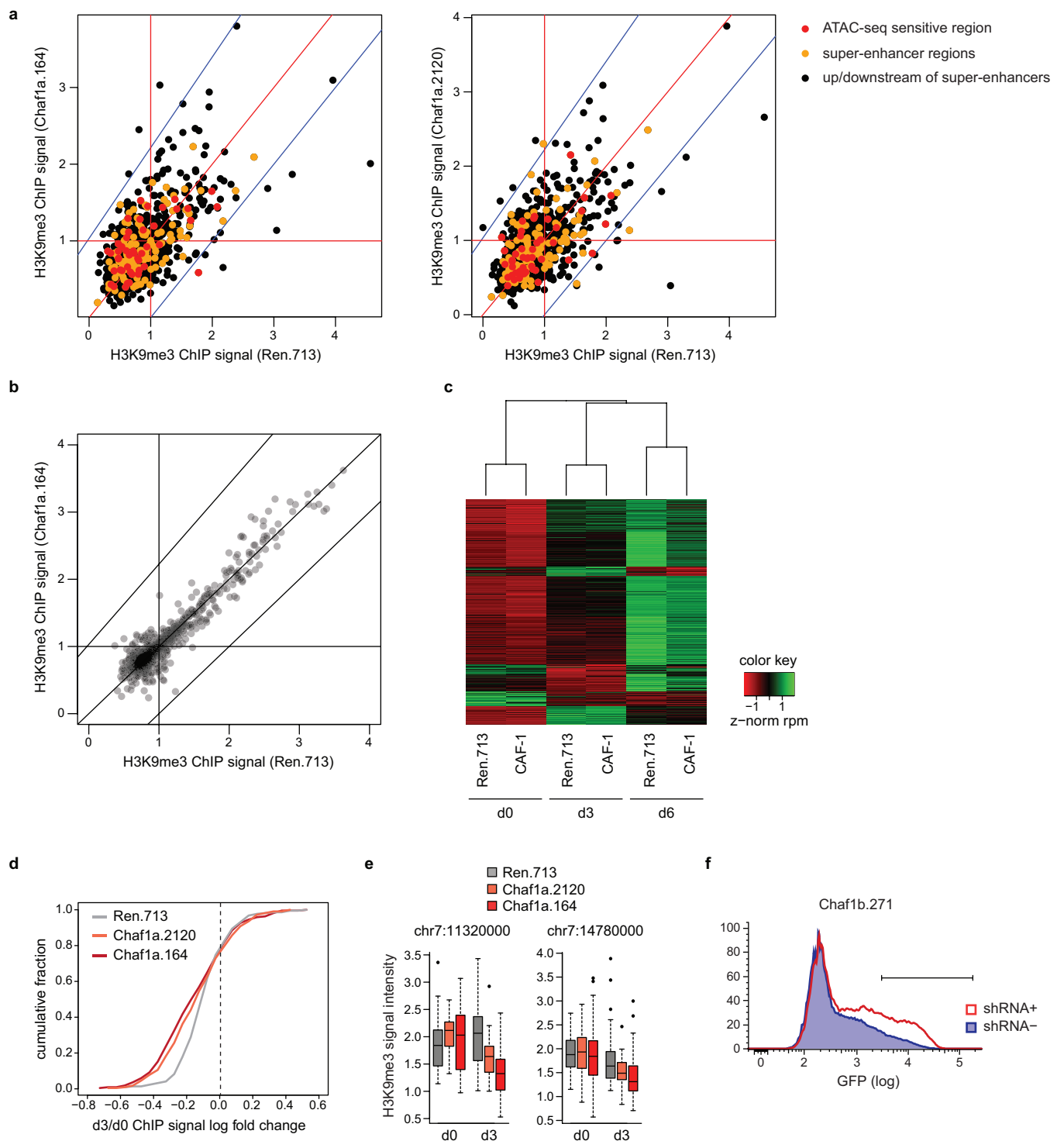
Chaf1a and *Renilla* control cells at day 3 of reprogramming to measure global chromatin accessibility over pluripotency-specific super-enhancer elements. ATAC-seq data from Chaf1a.164 shRNA- and Chaf1a.2120 shRNA-transduced cells were merged for this analysis. **e**, ATAC-seq accessibility maps at super-enhancer elements associated with the *Sall4* locus. Shaded grey bars highlight more accessible sites in Chaf1a knockdown samples at days 3 and 6 of reprogramming compared to *Renilla* shRNA controls. **f**, ATAC-seq analysis of Chaf1a and control cells at day 3 of reprogramming to measure global chromatin accessibility over lineage-specific super-enhancer elements<sup>43</sup> (C2C12, myoblast cell line; proB, progenitor B cells; Th, T helper cells). The  $n$  denotes number of examined enhancer elements for each cell type. ATAC-seq data from Chaf1a.164 shRNA- and Chaf1a.2120 shRNA-transduced cells were merged for this analysis.



**Extended Data Figure 8 | CAF-1 suppression facilitates Sox2 binding to chromatin.** **a**, Sox2 ChIP-seq enrichment across pluripotency-specific super-enhancer elements at day 3 of reprogramming in the presence of indicated shRNA vectors. **b**, Venn diagram depicting shared and unique Sox2 targets in Chaf1a and *Renilla* knockdown cells. **c**, Bar graph shows the number and fraction of ES-cell-specific Sox2 targets (blue colour) among Sox2-bound sites that are unique to Chaf1a or *Renilla*

knockdown cells at day 3 of OKSM expression. **d**, Sox2 ChIP-seq analysis of Chaf1a and control shRNA-infected cells at day 3 of reprogramming to determine enrichment of Sox2 binding across lineage-specific super-enhancer elements<sup>43</sup> (C2C12, myoblast cell line; proB, progenitor B cells; Th, T helper cells;  $P$  value  $< 10^{-15}$  for all comparisons between Chaf1a knockdown cells and control).





**Extended Data Figure 9 | CAF-1 suppression induces specific depletion of H3K9me3 at somatic heterochromatin domains.** **a**, Scatter plots comparing H3K9me3 enrichment nearby ATAC-seq sensitive and super-enhancer regions between control (Ren.713) and Chaf1a knockdown cells (Chaf1a.164 and Chaf1a.2122) at day 3 of reprogramming. Values reflect normalized H3K9me3 ChIP signal (IP/input) for 5-kb genomic regions overlapping ATAC-seq sensitive regions (red), super-enhancer regions (orange) and regions within 50-kb upstream and downstream of super-enhancers (black). **b**, Scatter plots comparing H3K9me3 enrichment over transposable element (TE) families in control and Chaf1a knockdown cells at day 3 of reprogramming. Values reflect normalized H3K9me3 ChIP-seq signal (IP/input) over families of TEs in the mouse genome. **c**, Heatmap shows the relative changes (z-normalized) of TE family expression as

estimated by RNA sequencing in control and Chaf1a knockdown cells at day 0, 3 and 6 of reprogramming. Data are clustered using the *k*-means algorithm. **d**, Cumulative histogram showing the relative fraction of reprogramming-resistant regions (RRRs)<sup>29</sup> (x axis) that display negative or positive enrichment (fold change) of average H3K9me3 signal at day 3 of reprogramming in control and Chaf1a knockdown cells. Note that more RRRs exhibit depletion of H3K9me3 in Chaf1a knockdown samples. **e**, H3K9me3 ChIP-seq analysis of RRRs after 0 and 3 days of reprogramming. Box plots depict representative RRRs on chromosome 7 ( $P < 0.05$  for both shRNAs). See also Fig. 5d. **f**, Histogram plot showing activation of UAS-Oct4-GFP transgene upon suppression of Chaf1b (shRNA<sup>+</sup> line) in the presence of Gal4-VP16 fusion protein. See Fig. 5f for quantification.

# Immune homeostasis enforced by co-localized effector and regulatory T cells

Zhiduo Liu<sup>1</sup>, Michael Y. Gerner<sup>1</sup>, Nicholas Van Panhuys<sup>1</sup>, Andrew G. Levine<sup>2,3</sup>, Alexander Y. Rudensky<sup>2,3</sup> & Ronald N. Germain<sup>1</sup>

FOXP3<sup>+</sup> regulatory T cells (T<sub>reg</sub> cells) prevent autoimmunity by limiting the effector activity of T cells that have escaped thymic negative selection or peripheral inactivation. Despite the information available about molecular factors mediating the suppressive function of T<sub>reg</sub> cells, the relevant cellular events in intact tissues remain largely unexplored, and whether T<sub>reg</sub> cells prevent activation of self-specific T cells or primarily limit damage from such cells has not been determined. Here we use multiplex, quantitative imaging in mice to show that, within secondary lymphoid tissues, highly suppressive T<sub>reg</sub> cells expressing phosphorylated STAT5 exist in discrete clusters with rare IL-2-positive T cells that are activated by self-antigens. This local IL-2 induction of STAT5 phosphorylation in T<sub>reg</sub> cells is part of a feedback circuit that limits further autoimmune responses. Inducible ablation of T cell receptor expression by T<sub>reg</sub> cells reduces their regulatory capacity and disrupts their localization in clusters, resulting in uncontrolled effector T cell responses. Our data thus reveal that autoreactive T cells are activated to cytokine production on a regular basis, with physically co-clustering T cell receptor-stimulated T<sub>reg</sub> cells responding in a negative feedback manner to suppress incipient autoimmunity and maintain immune homeostasis.

T<sub>reg</sub> cells play crucial roles in maintaining immune homeostasis<sup>1–3</sup> and many molecular mechanisms contribute to the ability of these cells to interfere with damaging anti-self responses<sup>4,5</sup>. To explore how T<sub>reg</sub> cells are organized in secondary lymphoid tissues under homeostatic conditions<sup>6</sup>, we used a new method for high-resolution, multiplex examination of tissue sections termed ‘histo-cytometry’<sup>7–9</sup>. This technique permits quantitative, spatially resolved phenotyping of cells in tissue sections akin to analysis by flow cytometry, while also permitting measurement of activation state using anti-phosphopeptide reagents, and functional state using anti-cytokine antibodies.

## pSTAT5<sup>+</sup> T<sub>reg</sub> cells exist as discrete clusters

We took advantage of observations showing that interleukin-2 (IL-2) is indispensable for maintaining T<sub>reg</sub> function *in vivo*<sup>10–12</sup> by searching for phosphorylated STAT5 (pSTAT5)-positive T cells in sections from mouse lymph nodes. In the steady state, we could detect pSTAT5 signals using flow cytometry primarily in a fraction of FOXP3<sup>+</sup> T<sub>reg</sub> cells (Extended Data Fig. 1a), as reported<sup>13</sup>. The pSTAT5 signals in T<sub>reg</sub> cells were stringently dependent on IL-2, as they were eliminated after treatment with an IL-2-blocking antibody (Extended Data Fig. 1a, b) and in IL-2- but not IL-15-knockout mice (Extended Data Fig. 1c). Using FOXP3-eGFP reporter mice<sup>14</sup>, we observed GFP<sup>+</sup> T<sub>reg</sub> cells distributed evenly in the T cell zone, whereas the pSTAT5<sup>+</sup> T<sub>reg</sub> subset was primarily localized to a few discrete aggregates in the outer paracortical T cell region (Fig. 1a). Three-dimensional analysis of 350-μm tissue sections clearly showed that a variable number of pSTAT5<sup>+</sup> T<sub>reg</sub> cells formed a small group with tightly associated cells in the centre, which we term a cluster (Supplementary Video 1); this contrasted with the scattered pSTAT5<sup>+</sup> T<sub>reg</sub> cells described previously in spleen<sup>13</sup>. To quantify the imaging data, pSTAT5<sup>+</sup> T<sub>reg</sub> cells were selectively visualized (Fig. 1b, left) with coordinates corresponding to their positions in the tissue (Fig. 1b, middle). The data were further transformed into a contour plot based on cell densities (Fig. 1b, right). pSTAT5<sup>+</sup> T<sub>reg</sub> clusters were then

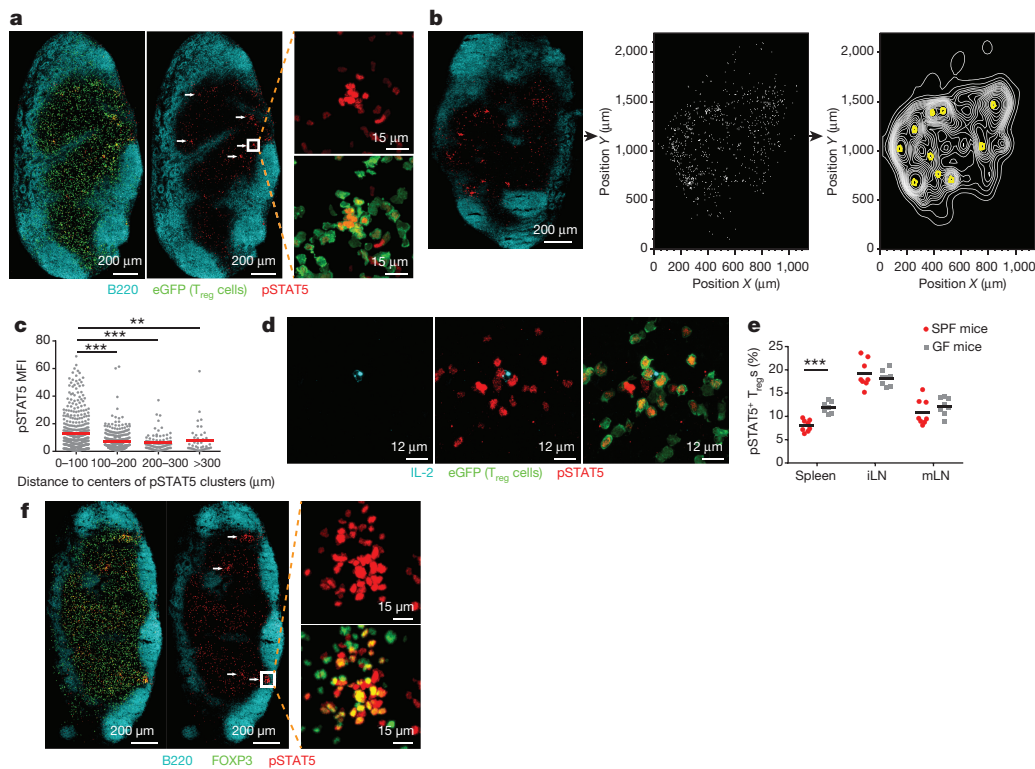
identified and the centre of each cluster determined. In addition to the uneven distribution of pSTAT5<sup>+</sup> T<sub>reg</sub> cells, their pSTAT5 signal intensity was heterogeneous. T<sub>reg</sub> cells less than 100 μm from the cluster centre showed significantly higher pSTAT5 intensity than did cells beyond that distance (Fig. 1c). This strongly suggested that the source of IL-2 driving pSTAT5 generation was in the cluster centre, and indeed, we detected individual IL-2 producing cells surrounded by pSTAT5<sup>+</sup> T<sub>reg</sub> cells (Fig. 1d) in many, but not all, such clusters. Among all the IL-2 producing cells we detected *in situ*, 75% were associated with T<sub>reg</sub> clusters (Extended Data Fig. 2); this is probably an underestimate given the sensitivity limitations of the imaging method and the temporal asynchrony of the responses. These *in vivo* data on the limited distance of strong pSTAT5 signals with respect to the cytokine producer cell agree with recent mathematical models of IL-2 signal spread in tissues<sup>15</sup>.

To identify the cells making IL-2 in the steady state we crossed Rag1<sup>-/-</sup> Il2<sup>GFP/GFP</sup> knock-in mice<sup>16</sup> with wild-type mice to generate Il2<sup>WT/GFP</sup> heterozygous mice, in which GFP expression from one allele allows for detection of IL-2 transcription, whereas the other cytokine-producing allele can protect the host from autoimmunity. The frequency of IL-2 producing cells in lymphoid tissues was extremely low (Extended Data Fig. 3a). Almost all the GFP<sup>+</sup> cells expressed CD3 but were negative for CD1d-tetramer staining, indicating they were not natural killer T cells. More than 80% of GFP<sup>+</sup>CD3<sup>+</sup> cells also expressed CD4, while fewer than 10% were CD8-positive (Extended Data Fig. 3b). These data indicate that CD4<sup>+</sup> T cells are the major cell type producing IL-2 in the steady state, as reported<sup>17–19</sup>.

T cells make IL-2 when the T cell receptor (TCR) stimulus reaches a certain threshold in the presence of co-stimulatory signals<sup>20</sup>. Previous reports of CD4<sup>+</sup> T cell production of IL-2 in the steady-state<sup>17–19</sup> did not examine whether the stimulus came from commensal organisms, the largest source of non-self-antigens presented by activated dendritic cells, or involved self-antigens presented by cells with adequate TCR ligand density and co-stimulatory properties even in the absence of

<sup>1</sup>Lymphocyte Biology Section, Laboratory of Systems Biology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892-1892, USA.

<sup>2</sup>Howard Hughes Medical Institute, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. <sup>3</sup>Immunology Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA.



**Figure 1 | pSTAT5<sup>+</sup> T<sub>reg</sub> clusters in lymph nodes.** **a**, Immunofluorescence staining of an inguinal lymph node section from SPF FOXP3-eGFP mice. Arrows indicate representative pSTAT5<sup>+</sup> T<sub>reg</sub> clusters. Images are representative of six lymph nodes total from six mice. **b**, Transformation of T<sub>reg</sub> image data into a contour plot. **c**, Correlation between T<sub>reg</sub> pSTAT5 intensity and distance to the nearest centre of a pSTAT5<sup>+</sup> cluster. Each dot indicates a pSTAT5<sup>+</sup> T<sub>reg</sub> cell in **b**. MFI, mean fluorescence intensity. **d**, Localization of an IL-2-producing cell within a pSTAT5<sup>+</sup> T<sub>reg</sub> cluster. **e**, Flow cytometry analysis of pSTAT5<sup>+</sup> T<sub>reg</sub> cells in different lymphoid

organs from SPF and germ-free (GF) mice. iLN, inguinal lymph node; mLN, mesenteric lymph node. Each dot indicates a single mouse. Results are pooled from three independent experiments.  $n = 8$  for all groups except GF spleen, for which  $n = 7$ . **f**, Immunofluorescence staining of an inguinal lymph node section from germ-free mice. Arrows indicate representative pSTAT5<sup>+</sup> T<sub>reg</sub> clusters. Images are representative of five lymph nodes total from five mice.  $**P < 0.01$ ,  $***P < 0.001$ , as calculated by one-way ANOVA with Tukey's post hoc test (**c**) or two-tailed Student's *t*-test (**e**).

microbial colonization. We therefore compared the number and nature of pSTAT5<sup>+</sup> T<sub>reg</sub> clusters in the lymphoid tissues of conventional specific-pathogen-free (SPF) versus germ-free animals. Remarkably, the lymphoid tissues of germ-free mice showed as many (and in the spleen, more) pSTAT5<sup>+</sup> T<sub>reg</sub> cells as compared to conventional animals (Fig. 1e). These clusters in germ-free animals were also similar in location and size to those of conventional SPF mice (Fig. 1f). This was true in skin-draining lymph nodes that would not be expected to be readily accessed by food-derived antigens, strongly suggesting that effector T cell activation for IL-2 production is driven by self-antigens. These latter findings indicate that even in the presence of functional T<sub>reg</sub> cells and absence of overt autoimmunity or discrete microbial stimuli, a small fraction of T cells can still be sufficiently activated by self-antigens to become IL-2-secreting proto-effector T cells. Such activation probably requires active suppression by T<sub>reg</sub> cells to prevent tissue-damaging responses, consistent with the evidence that T<sub>reg</sub> depletion in germ-free animals results in a severe lymphoproliferative disease and immune-mediated inflammation<sup>21</sup>.

### Identification of dendritic cells involved in T<sub>reg</sub> clustering

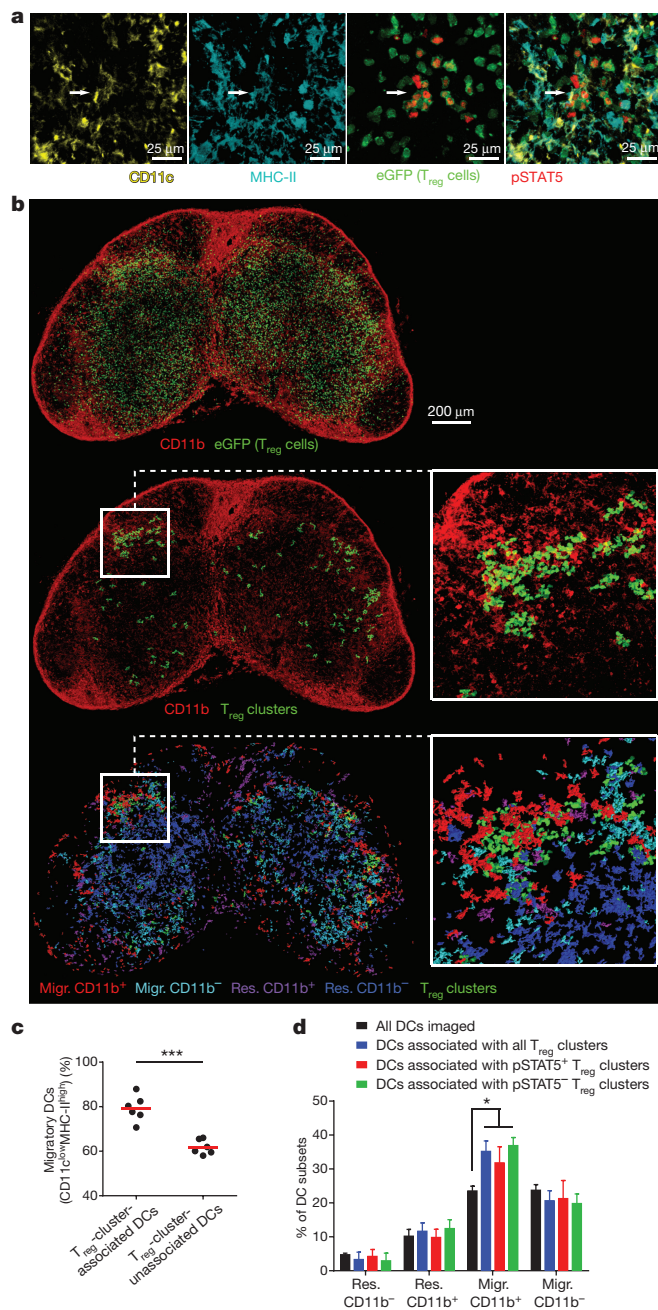
Based on prior observations using intravital two-photon microscopy showing migration arrest of T cells upon TCR engagement<sup>22,23</sup>, the tight clustering of T<sub>reg</sub> and effector T cells suggested that they were interacting with antigen-presenting cells, presumably dendritic cells, via self-ligand engagement by their TCR or following recent TCR-induced integrin affinity upregulation<sup>24</sup>. Indeed, CD11c<sup>+</sup> MHC class II (MHC-II)<sup>+</sup> dendritic cells could be directly observed in tight association with most T<sub>reg</sub> clusters (Fig. 2a). Given that the majority

of T<sub>reg</sub> clusters were localized in the outer T cell zone paracortex and interfollicular regions (Fig. 1a), our previous demonstration of differential dendritic cell subset positioning in steady-state lymph nodes<sup>7</sup> suggested that specific dendritic cell populations might be preferentially involved in cluster formation. Histo-cytometry showed that T<sub>reg</sub> clusters were predominantly localized in lymph node regions densely populated with CD11c<sup>low</sup>MHC-II<sup>high</sup> migratory dendritic cells (Extended Data Fig. 4a upper row, Fig. 2b). Quantitative analysis (Extended Data Fig. 4b) revealed a statistically significant preferential association of T<sub>reg</sub> clusters with migratory dendritic cells (Fig. 2c). Higher MHC-II and CD80/86 expression on such migratory cells is also likely to correlate with greater functional self-antigen presentation to T cells, thus making these dendritic cells excellent candidates for driving proto-effector activation. In accord with this notion, histo-cytometry indicated that there was a higher percentage of CD86<sup>high</sup> dendritic cells within T<sub>reg</sub> clusters than among dendritic cells not associated with such clusters (Extended Data Fig. 5). More detailed analysis showed that T<sub>reg</sub> cells were able to cluster with various dendritic cell subsets, with a preference for CD11b<sup>+</sup> migratory dendritic cells (Fig. 2d and Extended Data Fig. 4b). Together, these data show that in the steady state, self-reactive effector T cells co-aggregate with T<sub>reg</sub> cells around a subpopulation of mainly migratory dendritic cells presenting self-antigens.

### Suppressive phenotype of clustered T<sub>reg</sub> cells

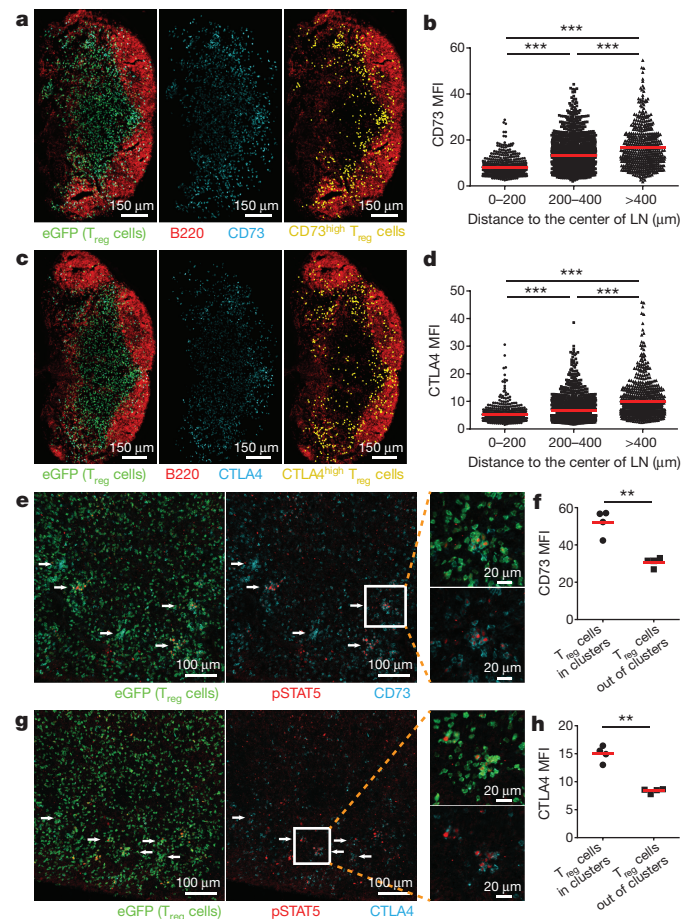
The known role of T<sub>reg</sub> cells in maintenance of immune homeostasis under steady-state conditions implies that there should be continuous successful suppression of the co-localized activated T cells to prevent an uncontrolled autoimmune response. T<sub>reg</sub> suppression of effector T cell responses has been reported to involve diverse molecular





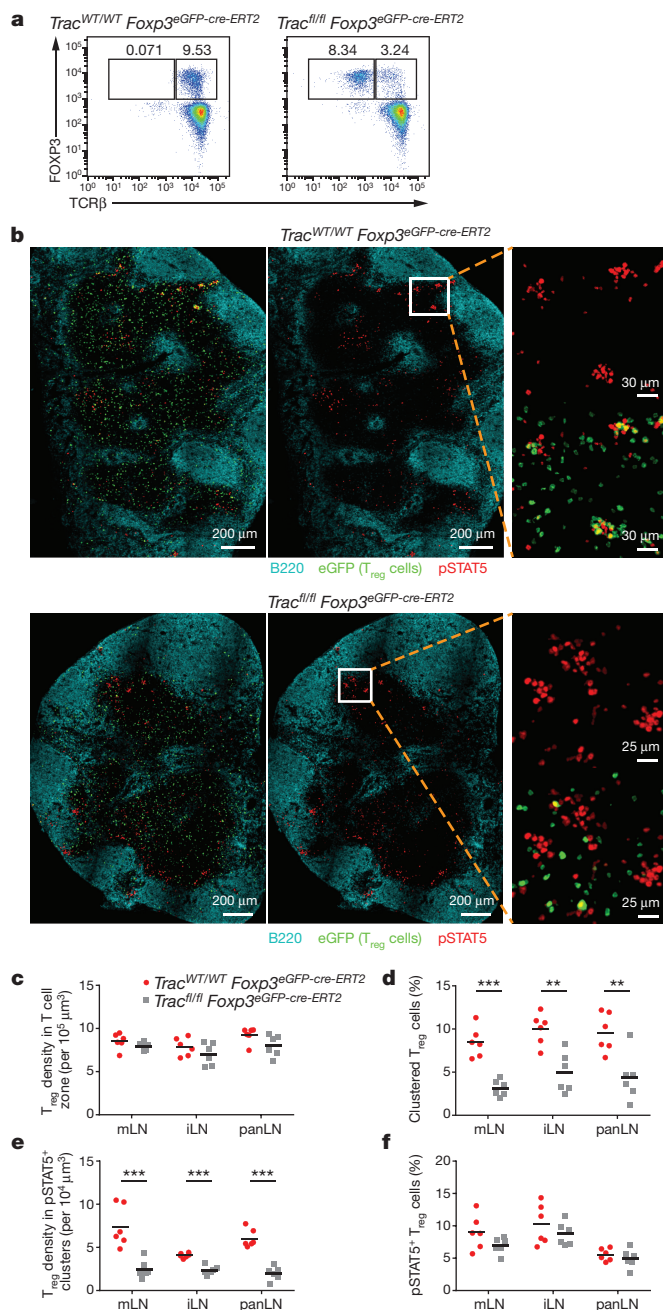
**Figure 2 | Preferential role of migratory dendritic cells in T<sub>reg</sub> clustering.** **a**, A representative image showing association of a T<sub>reg</sub> cluster with a CD11c<sup>+</sup>MHC-II<sup>+</sup> dendritic cell (arrow). **b**, Distribution of T<sub>reg</sub> clusters and different dendritic cell subsets in an inguinal lymph node section from a FOXP3-eGFP mouse. Creation of surfaces for each cell type was performed on the basis of immunofluorescence staining and subsequent histo-cytometry analysis as shown in Extended Data Fig. 4a. T<sub>reg</sub> clusters were determined by setting a threshold for the volume within T<sub>reg</sub> surfaces created in the eGFP channel without object splitting. Migr., migratory; Res., resident. **c**, Quantification of the percentages of migratory dendritic cells (DCs) associated with T<sub>reg</sub> clusters. Each dot indicates a section ( $n = 6$ , 6 sections total from 3 lymph nodes from 3 mice). **d**, Phenotypic characterization of dendritic cells associated with pSTAT5<sup>+</sup> and pSTAT5<sup>-</sup> T<sub>reg</sub> clusters ( $n = 6$ , 6 lymph nodes total from 6 mice, mean  $\pm$  s.e.m., pooled from two experiments). \* $P < 0.05$ , \*\*\* $P < 0.001$ , as calculated by two-tailed Student's *t*-test (c) or one-way ANOVA with Tukey's post hoc test (d).

mechanisms, including generation of adenosine via CD39/CD73 expression<sup>25,26</sup>, competition for IL-2 (ref. 27), and interference with co-stimulatory molecule engagement of CD28 on proto-effector T cells



**Figure 3 | Highly suppressive phenotype of clustered T<sub>reg</sub> cells.** **a**, Immunofluorescence staining of CD73 in a popliteal lymph node section from a FOXP3-eGFP mouse. CD73 signals were gated on eGFP<sup>+</sup> T<sub>reg</sub> cells. Yellow dots indicate the distribution of CD73<sup>high</sup> T<sub>reg</sub> cells. Images are representative of five lymph nodes total from five mice. **b**, The correlation between CD73 intensity of T<sub>reg</sub> cells and the distance of the T<sub>reg</sub> cells to the centre of the lymph node (LN). Each dot indicates a T<sub>reg</sub> cell shown in **a**. MFI, mean fluorescence intensity. **c**, Immunofluorescence staining of CTLA4 in a popliteal lymph node section from a FOXP3-eGFP mouse. CTLA4 signals were gated on eGFP<sup>+</sup> T<sub>reg</sub> cells. Yellow dots indicate the distribution of CTLA4<sup>high</sup> T<sub>reg</sub> cells. Images are representative of five lymph nodes total from five mice. **d**, The correlation between CTLA4 intensity of T<sub>reg</sub> cells and the distance of these T<sub>reg</sub> cells to the centre of the lymph node. Each dot indicates a T<sub>reg</sub> cell shown in **c**. **e**, Representative images showing CD73 staining in T<sub>reg</sub> clusters. Arrows indicate T<sub>reg</sub> clusters. **f**, Quantification of CD73 expression on T<sub>reg</sub> cells inside or outside of clusters ( $n = 4$ , 4 lymph nodes from 4 mice). Results are representative of two independent experiments. **g**, Representative images showing CTLA4 staining in T<sub>reg</sub> clusters. Arrows indicate T<sub>reg</sub> clusters. **h**, Quantification of CTLA4 expression on T<sub>reg</sub> cells inside or outside of clusters ( $n = 4$ , 4 lymph nodes from 4 mice). Results are representative of two independent experiments. \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , as calculated by one-way ANOVA with Tukey's post hoc test (b, d) or two-tailed Student's *t*-test (f, h).

through the expression of CTLA4 by the T<sub>reg</sub> cells<sup>28,29</sup>. Histo-cytometry revealed that the majority of CD73<sup>high</sup> T<sub>reg</sub> cells localized along the T-B border (Fig. 3a, b), an area also enriched for pSTAT5<sup>+</sup> T<sub>reg</sub> clusters and MHC-II<sup>high</sup> dendritic cells. Furthermore, CD73 expression level on T<sub>reg</sub> cells was negatively associated with their distance to the centre of cluster (Extended Data Fig. 6a, b) and clustered T<sub>reg</sub> cells showed the highest levels of CD73 amongst all lymph node T<sub>reg</sub> cells (Fig. 3e, f). CTLA4 expression showed a similar distribution pattern (Extended Data Fig. 6c, d and Fig. 3c, d, g, h). Thus, the highest expression of inhibitory molecules is on T<sub>reg</sub> cells within the clusters, consistent with these cells being the most adept at constraining effector T cell responses.



**Figure 4 | Role of T<sub>reg</sub> TCR expression in clustering and suppression.**  
**a**, Detection of TCRβ on CD4<sup>+</sup> lymph node T cells from *Trac*<sup>WT/WT</sup> *Foxp3*<sup>eGFP-cre-ERT2</sup> and *Trac*<sup>fl/fl</sup> *Foxp3*<sup>eGFP-cre-ERT2</sup> mice 7 days after tamoxifen treatment. Numbers adjacent to outlined areas indicate the percentages of cells in each area. Two mice for each group.  
**b**, Immunofluorescence staining of mesenteric lymph node sections from *Trac*<sup>WT/WT</sup> *Foxp3*<sup>eGFP-cre-ERT2</sup> (top) and *Trac*<sup>fl/fl</sup> *Foxp3*<sup>eGFP-cre-ERT2</sup> (bottom) mice treated as in **a**. **c–f**, Quantification of T<sub>reg</sub> density in T cell zone (**c**), percentage of T<sub>reg</sub> cells forming clusters (**d**), T<sub>reg</sub> density in pSTAT5<sup>+</sup> clusters (**e**), and percentage of pSTAT5<sup>+</sup> T<sub>reg</sub> cells (**f**). mLN, mesenteric lymph node; iLN, inguinal lymph node; panLN, pancreatic lymph node. Each dot indicates a section, and each type of lymph node includes six sections total from two lymph nodes from two individual mice. \*\*\*P < 0.001, \*\*P < 0.01, as calculated by two-tailed Student's *t*-test.

### Role of T<sub>reg</sub> TCR in clustering and suppression

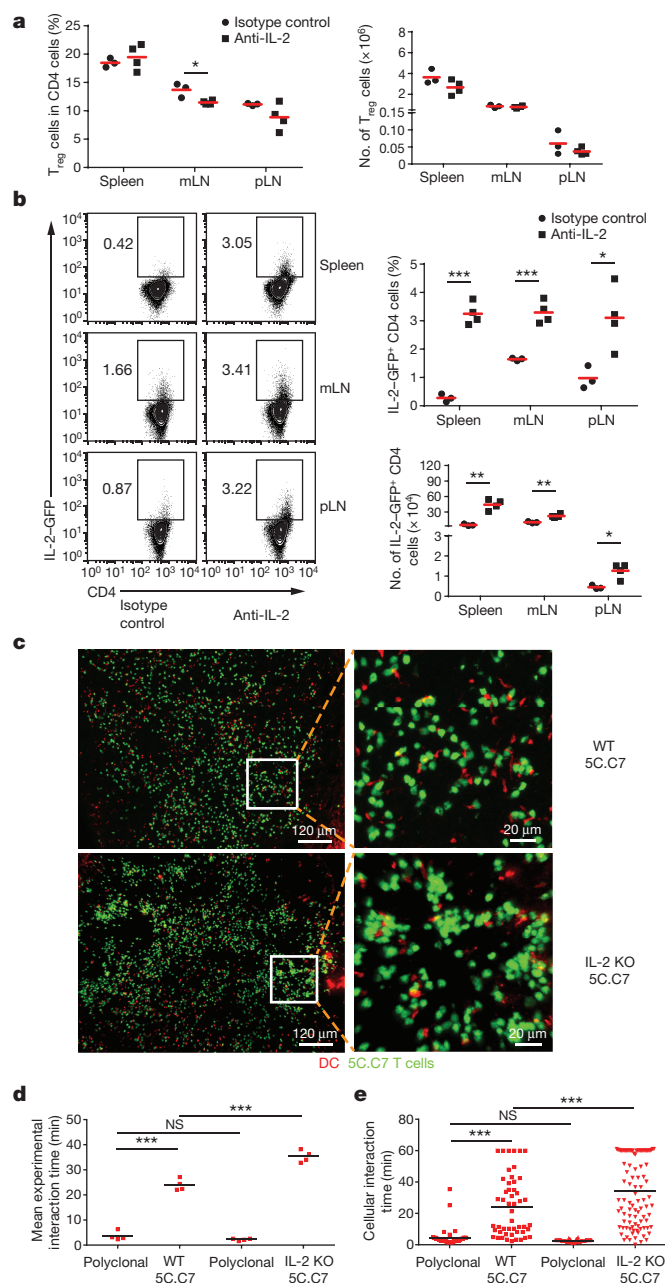
The tight conjugation of T<sub>reg</sub> cells to dendritic cells in the clusters implied that they were actively recognizing or had very recently recognized TCR ligands. Depriving mature T<sub>reg</sub> cells of TCR expression results in rampant autoimmunity within a few weeks without gross

diminution in the overall T<sub>reg</sub> number, indicating that post-thymic T<sub>reg</sub> cells need TCR expression to mediate their suppressive function<sup>30,31</sup>. Together, these data indicate that TCR engagement might contribute to effective T<sub>reg</sub> activity by participating in the co-localization of T<sub>reg</sub> cells with proto-effectors associated with stimulatory dendritic cells, as well as also possibly signalling the T<sub>reg</sub> to adopt a more suppressive state. To investigate this possibility, *Trac*<sup>fl/fl</sup> mice were crossed with *Foxp3*<sup>eGFP-cre-ERT2</sup> mice to enable tamoxifen-inducible deletion of TCRα specifically in T<sub>reg</sub> cells<sup>30</sup>. Mice were treated with tamoxifen on days 0 and 1, and euthanized for analysis on day 7. The majority of T<sub>reg</sub> cells (~70%) lost surface expression of TCRβ in *Trac*<sup>fl/fl</sup> *Foxp3*<sup>eGFP-cre-ERT2</sup> mice after tamoxifen treatment (Fig. 4a) but they remained in the T zone with unchanged overall density (Fig. 4b, c). However, the loss of TCR expression led to a striking difference in T<sub>reg</sub> spatial distribution of *Trac*<sup>WT/WT</sup> *Foxp3*<sup>eGFP-cre-ERT2</sup> versus *Trac*<sup>fl/fl</sup> *Foxp3*<sup>eGFP-cre-ERT2</sup> mice. The fraction of T<sub>reg</sub> cells in discrete clusters diminished in the lymph nodes of the treated *fl/fl* mice (Fig. 4b, d). Furthermore, T<sub>reg</sub> cells became a substantially smaller fraction of the cells in pSTAT5<sup>+</sup> clusters (Fig. 4b, e). The latter change was not accompanied by an overall loss of pSTAT5<sup>+</sup> T<sub>reg</sub> cells in the lymph nodes (Fig. 4f), in agreement with prior observations by flow cytometry<sup>30</sup>. Thus, acute loss of TCR signalling in T<sub>reg</sub> cells was coincident with a failure of these cells to form discrete clusters, with enhanced IL-2 production and signalling among proto-effector T cells, and with a loss of spatial constraint in IL-2 signalling in T<sub>reg</sub> cells. We then compared the expression of CD73 and CTLA4 by T<sub>reg</sub> cells of *Trac*<sup>WT/WT</sup> *Foxp3*<sup>eGFP-cre-ERT2</sup> versus *Trac*<sup>fl/fl</sup> *Foxp3*<sup>eGFP-cre-ERT2</sup> animals after tamoxifen treatment (Extended Data Fig. 7). The bulk population showed a small but significant change in mean CD73 (Extended Data Fig. 7a, b left) or CTLA4 (Extended Data Fig. 7c, d left) expression, and the highest-expressing cells among T<sub>reg</sub> cells were diminished substantially in the TCR-deleted animals. Furthermore, the high-expressing cells lost in the absence of T<sub>reg</sub> TCR engagement were those found preferentially in peripheral clusters in wild-type animals (Extended Data Fig. 7b, right and Extended Data Fig. 7d, right). These data on the selective reduction in T<sub>reg</sub> cells with the greatest capacity for suppression are consistent with the loss of T<sub>reg</sub> control of effector function in the tamoxifen-treated *fl/fl* mice<sup>30,31</sup>.

### IL-2 feedback regulates T<sub>reg</sub>-mediated suppression

In addition to depriving effector T cells of access to their own IL-2, which is important for effector T cell differentiation<sup>32</sup>, the advantage T<sub>reg</sub> cells have over conventional T cells<sup>33</sup> in the capture of IL-2 might play an important role in enhancing local T<sub>reg</sub> suppressive function. To test this hypothesis, we first blocked IL-2 in the steady state in *Il2*<sup>WT/GFP</sup> heterozygous mice by injecting IL-2-neutralizing antibody for 8 days, which did not grossly affect T<sub>reg</sub> numbers (Fig. 5a). However, there was a significant increase in both the frequency and the total number of IL-2/GFP<sup>+</sup> CD4<sup>+</sup> polyclonal T cells in spleen and in lymph nodes under these conditions (Fig. 5b). These findings are consistent with the idea that IL-2 could be playing a critical negative feedback role under normal conditions, informing T<sub>reg</sub> cells of incipient autoimmune responses of conventional T cells and by increasing T<sub>reg</sub> suppressive activity, constraining potentially damaging responses. To examine this possibility, we employed a cell transfer system involving pigeon cytochrome *c* (PCC)-specific 5C.C7 TCR transgenic T cells (5C.C7 T cells). PCC-peptide-pulsed splenic dendritic cells were transferred into wild-type mice possessing a full repertoire of T<sub>reg</sub> cells and normal endogenous levels of IL-2. Eighteen hours after dendritic cell transfer, alternately labelled polyclonal CD4<sup>+</sup> T cells and wild-type or IL-2-deficient 5C.C7 T cells were transferred into host animals. In this experimental system, endogenous T<sub>reg</sub> cells in the draining popliteal lymph node can sense locally generated IL-2 secreted by activated wild-type 5C.C7 T cells, but are blind to such cytokine feedback sensing with respect to the IL-2-knockout 5C.C7 T cells. Twenty-four hours after 5C.C7 T cell transfer, we could find large clusters of blasting IL-2-knockout 5C.C7 T cells binding to antigen-pulsed dendritic





**Figure 5 | Role of IL-2 feedback in  $T_{reg}$ -mediated suppression.**

**a**, Quantification of percentages and absolute numbers of  $T_{reg}$  cells in  $IL2^{WT/GFP}$  reporter mice treated with isotype control antibody ( $n=3$ ) or IL-2-neutralizing antibody ( $n=4$ ) for 8 days (1 mg per mouse daily). mLN, mesenteric lymph node; pLN, popliteal lymph node. **b**, Left panel, representative flow data showing  $GFP^{+}CD4^{+}$  cells in different lymphoid organs from  $IL2^{WT/GFP}$  reporter mice treated as in **a**. Right panel, quantification of percentages and absolute numbers of  $GFP^{+}CD4^{+}$ . Each dot in **a** and **b** indicates a single mouse. Results are representative of two independent experiments. **c**, Confocal imaging of popliteal lymph nodes 24 h after transfer with PCC-peptide-loaded splenic dendritic cells (DCs) and wild-type (WT) or  $IL2^{-/-}$  (IL-2 KO) 5C.C7 transgenic  $Rag2^{-/-}$  T cells. **d**, Quantification of interaction time between dendritic cells and wild-type or  $IL2^{-/-}$  5C.C7 transgenic  $Rag2^{-/-}$  T cells from four independent two-photon intravital imaging experiments. To control experimental preparation and to permit data comparison between different experiments, polyclonal T cells were co-transferred with each type of 5C.C7 cells. See also Supplementary Video 2. **e**, Cellular interaction time for individual cells pooled from four experiments. \*\*\* $P < 0.001$ , \*\* $P < 0.01$ , \* $P < 0.05$ , as calculated by two-tailed Student's  $t$ -test (**a**, **b**) or one-way ANOVA with Tukey's post hoc test (**d**, **e**). NS, not significant.

cells, whereas smaller 5C.C7 clusters were observed in lymph node of mice transferred with wild-type cells (Fig. 5c). In accord with this finding, intravital two-photon microscopy revealed a significantly longer interaction time between IL-2-knockout 5C.C7 T cells and dendritic cells (Supplementary video 2 and Fig. 5d, e) as compared to wild-type 5C.C7 T cells. Taken together, and in accord with a previous report<sup>34</sup>, these observations all support a critical role for IL-2-mediated feedback enhancement of  $T_{reg}$  suppressive function in controlling ongoing self-stimulation of proto-effector T cells and maintaining immune tolerance.

## Discussion

Our detection of  $pSTAT5^{+} T_{reg}$  cells that arise in response to IL-2 secreted by activated  $CD4^{+}$  T cells under steady-state conditions is inconsistent with the view that the immune system is kept in an 'off' state by tonic action of  $T_{reg}$  cells. Rather,  $T_{reg}$  cells appear to operate, at least in large measure, to constrain the expansion and development of conventional T cells into damaging effectors. There is a substantial risk difference between a system in which T cells only activate to self in the absence of  $T_{reg}$  cells and one in which the latter limit damage from already activated cells. These new findings showing the constant entry of auto-responsive cells into an activated state also account for the very rapid onset of autoimmune disease when  $T_{reg}$  cells are removed from the system or made ineffective by eliminating IL-2 feedback control or the capacity to act locally through TCR-dependent clustering<sup>30,31</sup>.

We also observed  $pSTAT5^{-} T_{reg}$  clusters in the lymph node, which may suggest either an absence of effector T cells in the cluster, or that activation and cytokine production by the proto-effector have already ceased. Both possibilities may exist in the steady state, as  $T_{reg}$  cells can either suppress the ongoing activation of effector T cells interacting with the same dendritic cell, or prevent some pre-effector T cells from binding to dendritic cells by maintaining dendritic cells in a less stimulatory state<sup>35,36</sup>.

Continuous TCR signalling plays a critical role in  $T_{reg}$  function in the periphery. Such signalling may act to maintain or induce the expression of genes whose products are required for effective suppression. Our data suggest that TCR signalling is probably also required for effective control of autoimmunity by promoting the co-localization of  $T_{reg}$  cells with target T effectors on a dendritic cell platform, although such clustering may only be optimized rather than solely mediated by a TCR-dependent mechanism. This process allows  $T_{reg}$  cells to stay together long enough with effector T cells to accomplish suppression, it maximizes the effects of inhibitory mechanisms, especially for membrane molecules and soluble factors with a limited working distance, and it promotes acquisition of enhanced suppressor capacity as evidenced by the high expression of CD73 and CTLA4 by clustered  $T_{reg}$  cells. These insights reveal the strategy used by  $T_{reg}$  cells to maintain effective tolerance, placing past observations on molecular mechanisms in an *in vivo* context that highlights the role of spatial proximity in this critical immunoregulatory process.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 19 April; accepted 2 November 2015.**

**Published online 25 November 2015.**

- Benoist, C. & Mathis, D.  $T_{reg}$  cells, life history, and diversity. *Cold Spring Harb. Perspect. Biol.* **4**, a007021 (2012).
- Josefowicz, S. Z., Lu, L. F. & Rudensky, A. Y. Regulatory T cells: mechanisms of differentiation and function. *Annu. Rev. Immunol.* **30**, 531–564 (2012).
- Xing, Y. & Hogquist, K. A. T-cell tolerance: central and peripheral. *Cold Spring Harb. Perspect. Biol.* **4**, (2012).
- Vignali, D. A., Collison, L. W. & Workman, C. J. How regulatory T cells work. *Nature Rev. Immunol.* **8**, 523–532 (2008).
- Shevach, E. M. Mechanisms of Foxp3<sup>+</sup> T regulatory cell-mediated suppression. *Immunity* **30**, 636–645 (2009).
- Germain, R. N. Maintaining system homeostasis: the third law of Newtonian immunology. *Nature Immunol.* **13**, 902–906 (2012).



7. Gerner, M. Y., Kastenmuller, W., Ifrim, I., Kabat, J. & Germain, R. N. Histo-cytometry: a method for highly multiplex quantitative tissue imaging analysis applied to dendritic cell subset microanatomy in lymph nodes. *Immunity* **37**, 364–376 (2012).
8. Gerner, M. Y., Torabi-Parizi, P. & Germain, R. N. Strategically localized dendritic cells promote rapid T cell responses to lymph-borne particulate antigens. *Immunity* **42**, 172–185 (2015).
9. Radtke, A. J. *et al.* Lymph-node resident CD8 $\alpha^+$  dendritic cells capture antigens from migratory malaria sporozoites and induce CD8 $^+$  T cell responses. *PLoS Pathog.* **11**, e1004637 (2015).
10. Fontenot, J. D., Rasmussen, J. P., Gavin, M. A. & Rudensky, A. Y. A function for interleukin 2 in Foxp3-expressing regulatory T cells. *Nature Immunol.* **6**, 1142–1151 (2005).
11. D'Cruz, L. M. & Klein, L. Development and function of agonist-induced CD25 $^+$ Foxp3 $^+$  regulatory T cells in the absence of interleukin 2 signaling. *Nature Immunol.* **6**, 1152–1159 (2005).
12. Cheng, G., Yu, A. & Malek, T. R. T-cell tolerance and the multi-functional role of IL-2R signaling in T-regulatory cells. *Immunol. Rev.* **241**, 63–76 (2011).
13. Smigielski, K. S. *et al.* CCR7 provides localized access to IL-2 and defines homeostatically distinct regulatory T cell subsets. *J. Exp. Med.* **211**, 121–136 (2014).
14. Bettelli, E. *et al.* Reciprocal developmental pathways for the generation of pathogenic effector TH17 and regulatory T cells. *Nature* **441**, 235–238 (2006).
15. Hofer, T., Krichevsky, O. & Altan-Bonnet, G. Competition for IL-2 between regulatory and effector T cells to chisel immune responses. *Front. Immunol.* **3**, 268 (2012).
16. Naramura, M., Hu, R. J. & Gu, H. Mice with a fluorescent marker for interleukin 2 gene activation. *Immunity* **9**, 209–216 (1998).
17. Setoguchi, R., Hori, S., Takahashi, T. & Sakaguchi, S. Homeostatic maintenance of natural Foxp3 $^+$  CD25 $^+$  CD4 $^+$  regulatory T cells by interleukin (IL)-2 and induction of autoimmune disease by IL-2 neutralization. *J. Exp. Med.* **201**, 723–735 (2005).
18. Almeida, A. R., Zaragoza, B. & Freitas, A. A. Indexation as a novel mechanism of lymphocyte homeostasis: the number of CD4 $^+$ CD25 $^+$  regulatory T cells is indexed to the number of IL-2-producing cells. *J. Immunol.* **177**, 192–200 (2006).
19. Amado, I. F. *et al.* IL-2 coordinates IL-2-producing and regulatory T cell interplay. *J. Exp. Med.* **210**, 2707–2720 (2013).
20. Powell, J. D., Ragheb, J. A., Kitagawa-Sakakida, S. & Schwartz, R. H. Molecular regulation of interleukin-2 expression by CD28 co-stimulation and anergy. *Immunol. Rev.* **165**, 287–300 (1998).
21. Chinen, T., Volchkov, P. Y., Chervonsky, A. V. & Rudensky, A. Y. A critical role for regulatory T cell-mediated control of inflammation in the absence of commensal microbiota. *J. Exp. Med.* **207**, 2323–2330 (2010).
22. Mempel, T. R., Henrickson, S. E. & Von Andrian, U. H. T-cell priming by dendritic cells in lymph nodes occurs in three distinct phases. *Nature* **427**, 154–159 (2004).
23. van Panhuys, N., Klauschen, F. & Germain, R. N. T-cell-receptor-dependent signal intensity dominantly controls CD4 $^+$  T cell polarization *in vivo*. *Immunity* **41**, 63–74 (2014).
24. Wang, H., Lim, D. & Rudd, C. E. Immunopathologies linked to integrin signalling. *Semin. Immunopathol.* **32**, 173–182 (2010).
25. Deaglio, S. *et al.* Adenosine generation catalysed by CD39 and CD73 expressed on regulatory T cells mediates immune suppression. *J. Exp. Med.* **204**, 1257–1265 (2007).
26. Kobia, J. J. *et al.* T regulatory and primed uncommitted CD4 T cells express CD73, which suppresses effector CD4 T cells by converting 5'-adenosine monophosphate to adenosine. *J. Immunol.* **177**, 6780–6786 (2006).
27. Pandiyan, P., Zheng, L., Ishihara, S., Reed, J. & Lenardo, M. J. CD4 $^+$ CD25 $^+$ Foxp3 $^+$  regulatory T cells induce cytokine deprivation-mediated apoptosis of effector CD4 $^+$  T cells. *Nature Immunol.* **8**, 1353–1362 (2007).
28. Read, S. *et al.* Blockade of CTLA-4 on CD4 $^+$ CD25 $^+$  regulatory T cells abrogates their function *in vivo*. *J. Immunol.* **177**, 4376–4383 (2006).
29. Wing, K. *et al.* CTLA-4 control over Foxp3 $^+$  regulatory T cell function. *Science* **322**, 271–275 (2008).
30. Levine, A. G., Arvey, A., Jin, W. & Rudensky, A. Y. Continuous requirement for the TCR in regulatory T cell function. *Nature Immunol.* **15**, 1070–1078 (2014).
31. Vahl, J. C. *et al.* Continuous T cell receptor signals maintain a functional regulatory T cell pool. *Immunity* **41**, 722–736 (2014).
32. Liao, W., Lin, J. X. & Leonard, W. J. Interleukin-2 at the crossroads of effector responses, tolerance, and immunotherapy. *Immunity* **38**, 13–25 (2013).
33. Lu, L. F. *et al.* Foxp3-dependent microRNA155 confers competitive fitness to regulatory T cells by targeting SOCS1 protein. *Immunity* **30**, 80–91 (2009).
34. O'Gorman, W. E. *et al.* The initial phase of an immune response functions to activate regulatory T cells. *J. Immunol.* **183**, 332–339 (2009).
35. Tang, Q. *et al.* Visualizing regulatory T cell control of autoimmune responses in nonobese diabetic mice. *Nature Immunol.* **7**, 83–92 (2006).
36. Tadokoro, C. E. *et al.* Regulatory T cells inhibit stable contacts between CD4 $^+$  T cells and dendritic cells *in vivo*. *J. Exp. Med.* **203**, 505–511 (2006).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank Y. Belkaid for providing germ-free mice. We also would like to thank members of the Lymphocyte Biology Section, Laboratory of Systems Biology for their helpful comments during the course of these studies and critical input during preparation of this manuscript. This work was supported by the Intramural Research Program of NIAID, NIH and by the US National Institutes of Health (R37AI034206 to A.Y.R.; T32GM007739 to A.G.L.), the Ludwig Cancer Center at Memorial Sloan-Kettering Cancer Center (A.Y.R.), and the Howard Hughes Medical Institute (A.Y.R.).

**Author Contributions** Z.L. designed and conducted most of the experiments, data analysis, and drafted the manuscript; M.Y.G. performed the analysis of DC phenotype and provided assistance with histo-cytometry studies; N.V.P. conducted 2P intravital imaging experiments; A.G.L. and A.Y.R. created *Trac<sup>wt/wt</sup>Foxp3<sup>eGFP-cre-ERT2</sup>* and *Trac<sup>fl/fl</sup>Foxp3<sup>eGFP-cre-ERT2</sup>* mice, generated tissue samples from tamoxifen treatment of these mice, and discussed data interpretation; R.N.G. helped design experiments, interpret data, and with input from all other authors, develop the final version of the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.N.G. ([rgermain@nih.gov](mailto:rgermain@nih.gov)).

## METHODS

**Mice.** C57BL/6 and *Il2*<sup>-/-</sup> mice were obtained from Taconic Laboratories. FOXP3-eGFP, B10.A CD45.2<sup>-</sup>, B10.A CD45.2<sup>+</sup> 5C.C7 TCR-transgenic *Rag2*<sup>-/-</sup>, B10.A CD45.2<sup>+</sup> 5C.C7 TCR-transgenic *Rag2*<sup>-/-</sup> *Il2*<sup>-/-</sup>, and *Il2*<sup>GFP/GFP</sup> knock-in *Rag1*<sup>-/-</sup> mice were obtained from Taconic Laboratories through a special NIAID contract. *Il2*<sup>GFP/GFP</sup> *Rag1*<sup>-/-</sup> mice were crossed to C57BL/6 mice to generate *Il2*<sup>WT/GFP</sup> heterozygous mice. *Il2*<sup>-/-</sup> mice were obtained from Jackson Laboratories. All the above mice were maintained in SPF conditions at an Association for Assessment and Accreditation of Laboratory Animal Care-accredited animal facility. Germ-free C57BL/6 mice were kindly provided by Y. Belkaid (National Institutes of Health). *Trac*<sup>WT/WT</sup> *Foxp3*<sup>eGFP-cre-ERT2</sup> and *Trac*<sup>fl/fl</sup> *Foxp3*<sup>eGFP-cre-ERT2</sup> mice were generated as previously described<sup>30</sup>. Six- to ten-week-old female mice were used for all experiments, and were randomly allocated into treatment groups. The investigators were not blinded to allocation during experiments and outcome assessment. For tamoxifen administration, mice were given oral gavage of 200 µl tamoxifen emulsion per treatment on days 0 and 1, and euthanized for analysis on day 7. All procedures were approved by the NIAID Animal Care and Use Committee (NIH).

**Flow cytometry.** Cells were isolated from lymphoid organs as indicated and stained with anti-CD3 (17A2; Biolegend), anti-CD4 (L3T4; Biolegend), anti-CD8 (53-6.7; Biolegend), anti-FOXP3 (150D; Biolegend), anti-TCRb (H57-597; Biolegend), anti-GFP (Life Technologies) and CD1d-tetramer (NIH Tetramer facility). For pSTAT5 staining, cells were fixed using BD Cytotfix Fixation Buffer (10 min at 37°C), permeabilized in BD Phosflow Perm Buffer III (30 min on ice) and stained with anti-pSTAT5 (47/Stat5; BD Biosciences). Flow cytometric data were collected on an LSR II (BD Biosciences) and analysed with FlowJo software (TreeStar).

**IL-2 neutralization.** For examining the role of IL-2 in pSTAT5 signals in T<sub>reg</sub> cells, mice were injected intravenously with 1 mg anti-IL2 (S4B6-1; BioXcell) or isotype control antibody (2A3; BioXcell) 24 h before sacrifice. For examining the role of IL-2 in T<sub>reg</sub> function, *Il2*<sup>WT/GFP</sup> mice were treated intraperitoneally with 1 mg anti-IL-2 or isotype control daily for 8 days before analysis.

**Immunofluorescence staining and confocal imaging.** Lymph nodes were harvested and fixed using fixation buffer containing 1% PFA for 12 h, then dehydrated in 30% sucrose before embedding in OCT freezing media (Sakura Finetek). 20-µm sections were prepared with a Leica cryostat and blocked with a blocking buffer containing 1% normal mouse serum, 1% bovine serum albumin and 0.3% Triton X-100 for 1 h. Sections were stained with directly conjugated antibodies or appropriate primary and secondary antibodies for 3 h at room temperature or overnight at 4°C in a dark, humidified chamber.

For whole-mount staining, 350-µm lymph node sections were made with Leica VT1000 S Vibrating blade microtome and blocked for 8 h, followed by staining with antibodies diluted in the blocking buffer for 48 h at 4°C on a shaker. Tissue clearing was performed as previously detailed<sup>37</sup>.

The following antibodies were used for staining: anti-B220 (RA3-6B2; Biolegend), anti-CD3 (17A2; Biolegend), anti-CD8 (53-6.7; Biolegend), anti-FOXP3 (FJK-16 s; eBioscience), anti-IL-2 (JES6-5H4; eBioscience), anti-GFP (Life Technologies), anti-CD11c (N418; Biolegend), anti-MHC-II (M5/114.15.2; Biolegend), anti-CTLA4 (UC10-4B9; Biolegend), anti-CD73 (TY/11.8; Biolegend), anti-CD11b (5C6; AbD Serotec), anti-CD86 (AP-MAB0803; Novus Biologicals) and anti-pSTAT5 (C11C5; Cell Signaling Technology).

Images were acquired on a Zeiss 710 confocal microscope or a Leica Sp8 confocal microscope.

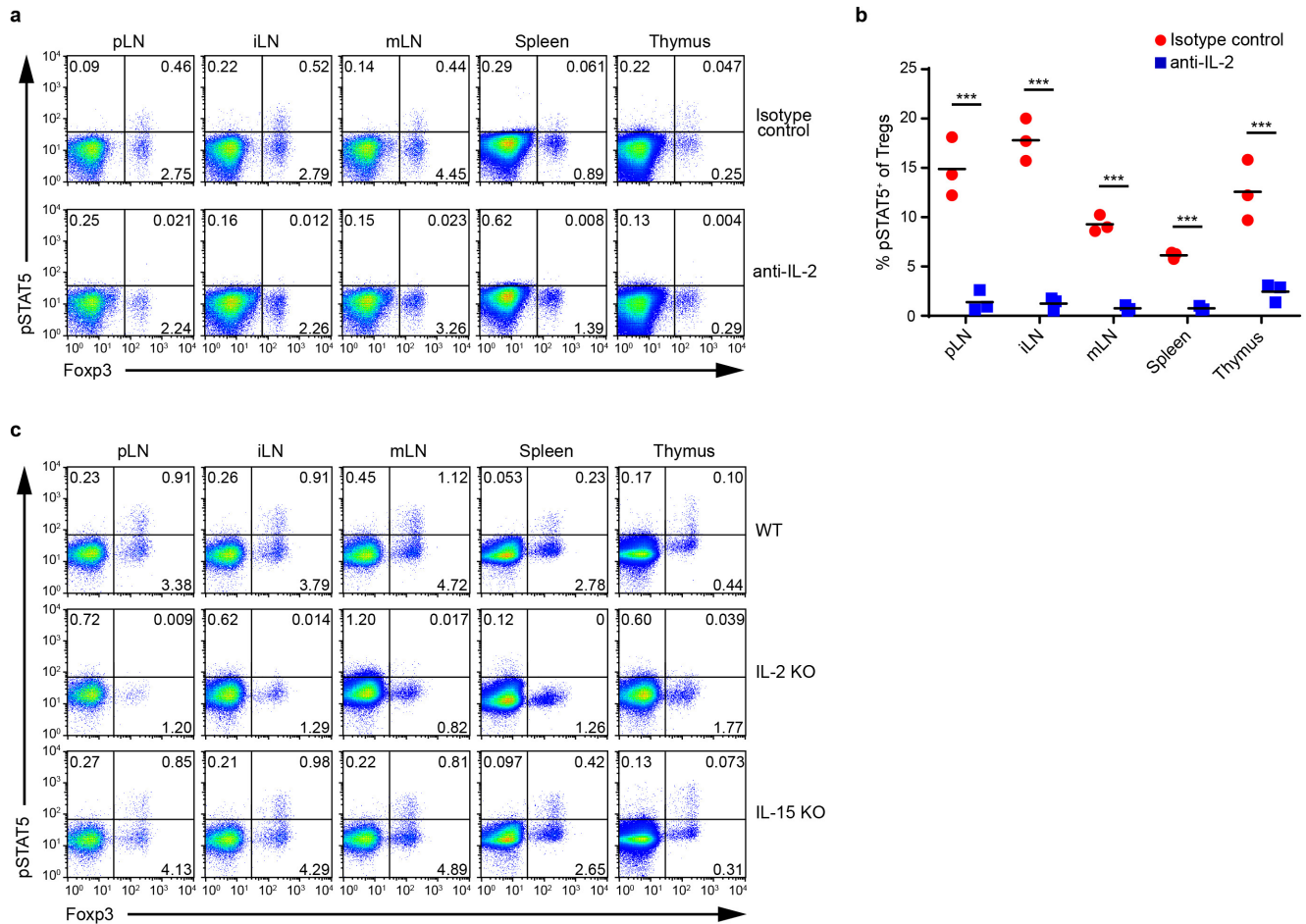
**Adoptive cell transfer and two-photon intravital imaging.** Cell preparation and adoptive transfer were performed as previously described<sup>23</sup>. In brief, splenic dendritic cells were purified from B10.A CD45.2<sup>-</sup> mice using CD11c microbeads (Miltenyi Biotec). Polyclonal CD4<sup>+</sup> T cells from lymph nodes of B10.A CD45.2<sup>-</sup> and TCR transgenic 5C.C7 CD4<sup>+</sup> T cells from wild-type or *Il2*<sup>-/-</sup> B10.A CD45.2<sup>+</sup> 5C.C7 TCR-transgenic *Rag2*<sup>-/-</sup> mice were purified by negative immunomagnetic cell sorting (Miltenyi Biotec). Dendritic cells were incubated *in vitro* with PCC peptide (10 µM pPCC, American Peptide Company) and LPS (1.0 µg ml<sup>-1</sup>, InvivoGen) for 4 h at 37°C before subcutaneous injection at 1 × 10<sup>6</sup> per footpad. CD4<sup>+</sup> T cells were transferred by intravenous injection at 2 × 10<sup>6</sup> per recipient 18 h post-transfer of dendritic cells.

For two-photon intravital imaging, dendritic cells were stained with 100 µM CTB (7-amino-4-chloromethylcoumarin, Molecular Probes), polyclonal CD4<sup>+</sup> T cells were stained with 1.25 µM CMFDA (5-chloromethylfluorescein diacetate, Molecular Probes), and wild-type or *Il2*<sup>-/-</sup> TCR transgenic 5C.C7 CD4<sup>+</sup> T cells were stained with 1.25 µM CMTX (Molecular Probes). Twenty-four hours after T cell transfer, mice were anaesthetized with isoflurane and popliteal lymph nodes were surgically exposed. Imaging was performed on a Zeiss 710 microscope equipped with a Chameleon laser (Coherent) tuned to 800 nm in combination with a 20× water-dipping lens (NA 1.0, Zeiss) using Zen 2010 acquisition software.

**Histo-cytometry.** Histo-cytometry analysis was performed as previously described<sup>7-9</sup>, with minor modifications. In brief, multi-parameter confocal images were corrected for fluorophore spillover using the Leica Channel Dye Separation module. Owing to the high spatial resolution of the 63 × 1.4 NA objective, deconvolution was not performed. For analysis of dendritic cell subsets associated with T<sub>reg</sub> clusters, all lymph node regions with visible T<sub>reg</sub> cell clusters were first imaged, with individual files then recombined into a single composite file representing each lymph node. To identify T<sub>reg</sub> clusters, the FOXP3-eGFP channel was used for T<sub>reg</sub> surface creation with zero object splitting (Imaris, Bitplane). T<sub>reg</sub> surfaces with a volume above a certain threshold were considered as T<sub>reg</sub> clusters. These T<sub>reg</sub> clusters were then separated based on pSTAT5 mean intensity parameter to isolate discrete pSTAT5<sup>+</sup> and pSTAT5<sup>-</sup> T<sub>reg</sub> clusters, which were then used to create new binary pSTAT5<sup>+</sup> and pSTAT5<sup>-</sup> T<sub>reg</sub> cluster channels. Dendritic cell surface marker gating within CD11c<sup>+</sup> MHC-II<sup>+</sup> CD3<sup>-</sup> B220<sup>-</sup> voxels was then performed as previously described<sup>7</sup>. Dendritic cell surfaces were next created based on a newly generated dendritic cell channel (CD11c<sup>+</sup> MHC-II<sup>+</sup> CD3<sup>-</sup> B220<sup>-</sup>), with the object statistics exported into FlowJo X (TreeStar Inc.) for analysis and graphing (Prism, Graphpad). Finally, dendritic cells that associate with T<sub>reg</sub> clusters were determined by gating on dendritic cell surfaces positive for intensity in the already created pSTAT5<sup>+</sup> or pSTAT5<sup>-</sup> T<sub>reg</sub> cluster channels.

**Statistical analysis.** Statistical analyses were performed using GraphPad Prism 6 software. Statistical tests were selected based on appropriate assumptions with respect to data distribution and variance characteristics. One-way ANOVA with Tukey's post hoc test was used for the statistical analysis of multiple groups. Student's *t*-test (two-tailed) was used for the statistical analysis of differences between two groups. No statistical methods were used to predetermine sample size.

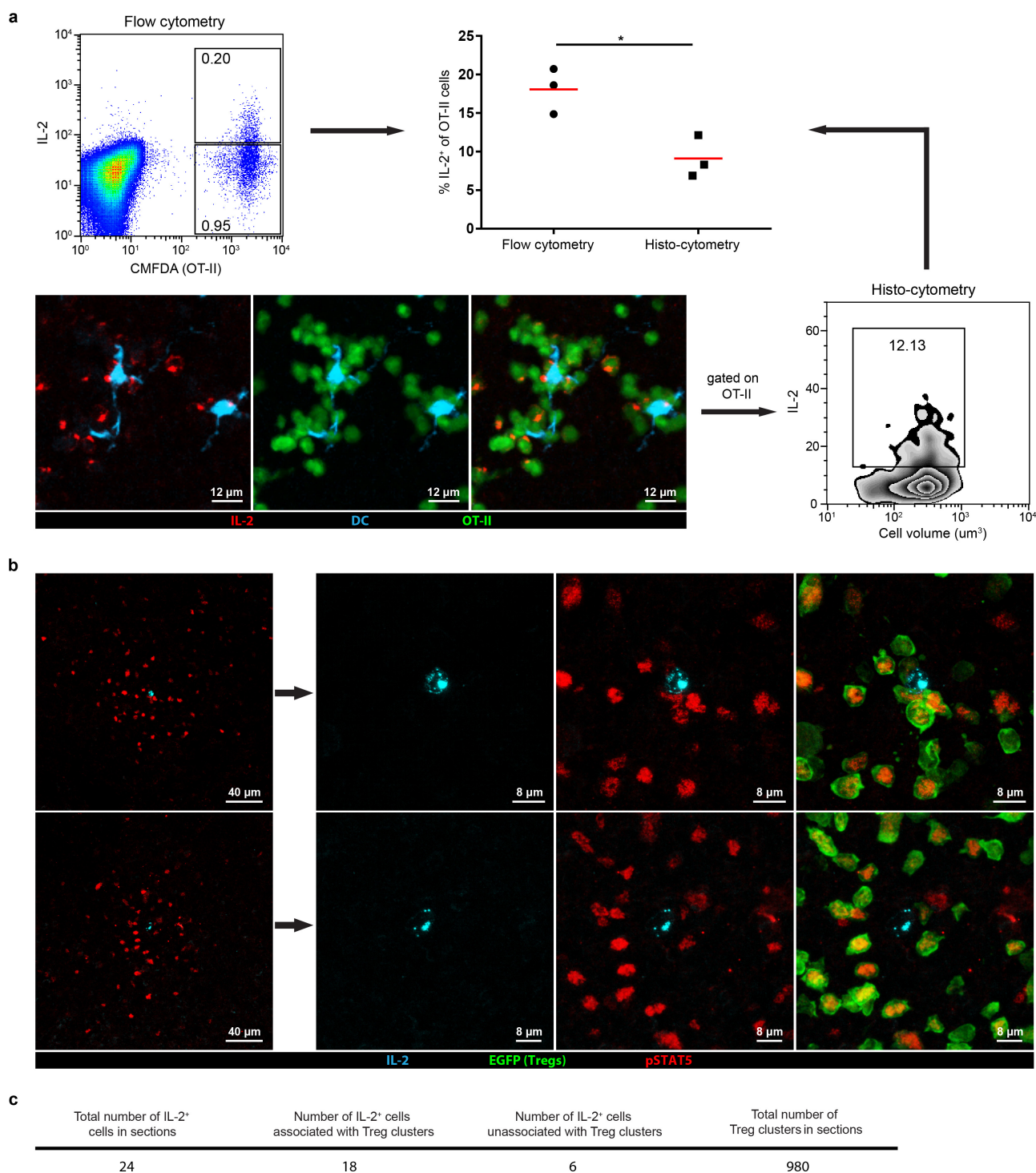
37. Ertürk, A. *et al.* Three-dimensional imaging of solvent-cleared organs using 3DISCO. *Nature Protocols* **7**, 1983–1995 (2012).



**Extended Data Figure 1 | Role of IL-2 in induction of pSTAT5 in  $T_{reg}$  cells.** **a**, Representative flow data showing pSTAT5 staining of  $T_{reg}$  cells from lymphoid organs as indicated after treatment with isotype antibody or IL-2-neutralizing antibody (1 mg per mouse) for 24 h. The numbers in the quadrants indicate the percentages of cells in each quadrant. pLN, popliteal lymph node; iLN, inguinal lymph node; mLN, mesenteric lymph

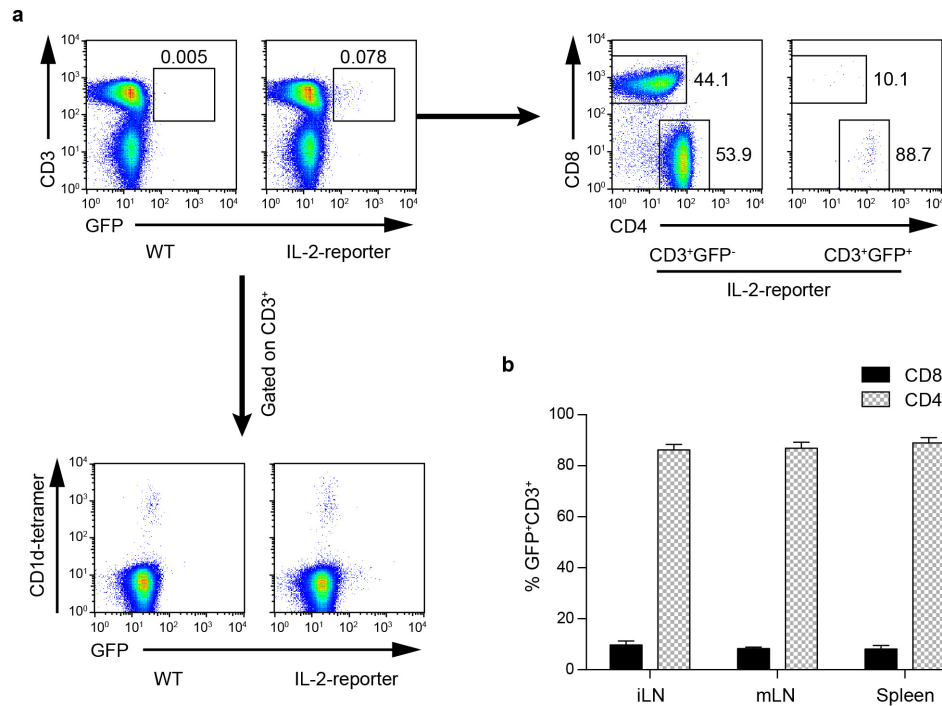
node. **b**, Quantification of pSTAT5<sup>+</sup>  $T_{reg}$  cells ( $n = 3$  for each treatment). Results are representative of two independent experiments. **c**, Detection of pSTAT5<sup>+</sup>  $T_{reg}$  cells in wild-type, IL-2-knockout and IL-15-knockout mice. The numbers in the quadrants indicate the percentages of cells in each quadrant. Results are representative of two independent experiments. \*\*\* $P < 0.001$ , as calculated by two-tailed Student's  $t$ -test.





**Extended Data Figure 2 | Spatial correlation between IL-2<sup>+</sup> cells and pSTAT5<sup>+</sup> T<sub>reg</sub> clusters.** **a**, A positive control showing immunofluorescence staining of IL-2 *in situ*, and the comparison of IL-2 detection sensitivity between flow cytometry and histo-cytometry. OVA<sub>323-339</sub> peptide-loaded splenic dendritic cells (DC, cyan) were stained with CMTMR (5-(and-6)-(((4-chloromethyl)benzoyl)amino)tetramethylrhodamine) and injected into the footpad, and Rag2<sup>-/-</sup> TCR transgenic OT-II CD4<sup>+</sup> T cells (green) stained with CMFDA (5-chloromethylfluorescein diacetate) were transferred 18 h post-transfer of dendritic cells. For each recipient, 16 h after T cell transfer, one draining popliteal lymph node was isolated for IL-2 (red) immunofluorescence staining, while the contralateral

lymph node was isolated and dissociated into single cells for intracellular IL-2 staining *ex vivo*. Each dot in the middle panel of the upper row indicates a lymph node ( $n = 3$ , 3 lymph nodes total from 3 mice for each analysis). \* $P < 0.05$  as calculated by two-tailed Student's *t*-test. **b**, Immunofluorescence staining of IL-2 and pSTAT5 in lymph nodes from FOXP3-eGFP reporter mice. Upper row indicates an IL-2<sup>+</sup> cell associated with clustered T<sub>reg</sub> cells among a group of pSTAT5<sup>+</sup> T<sub>reg</sub> cells. Lower row indicates an unassociated IL-2<sup>+</sup> cell surrounded by a group of pSTAT5<sup>+</sup> T<sub>reg</sub> cells. **c**, Quantification of IL-2<sup>+</sup> cells associated or unassociated with T<sub>reg</sub> clusters. Data are pooled from seven lymph nodes.

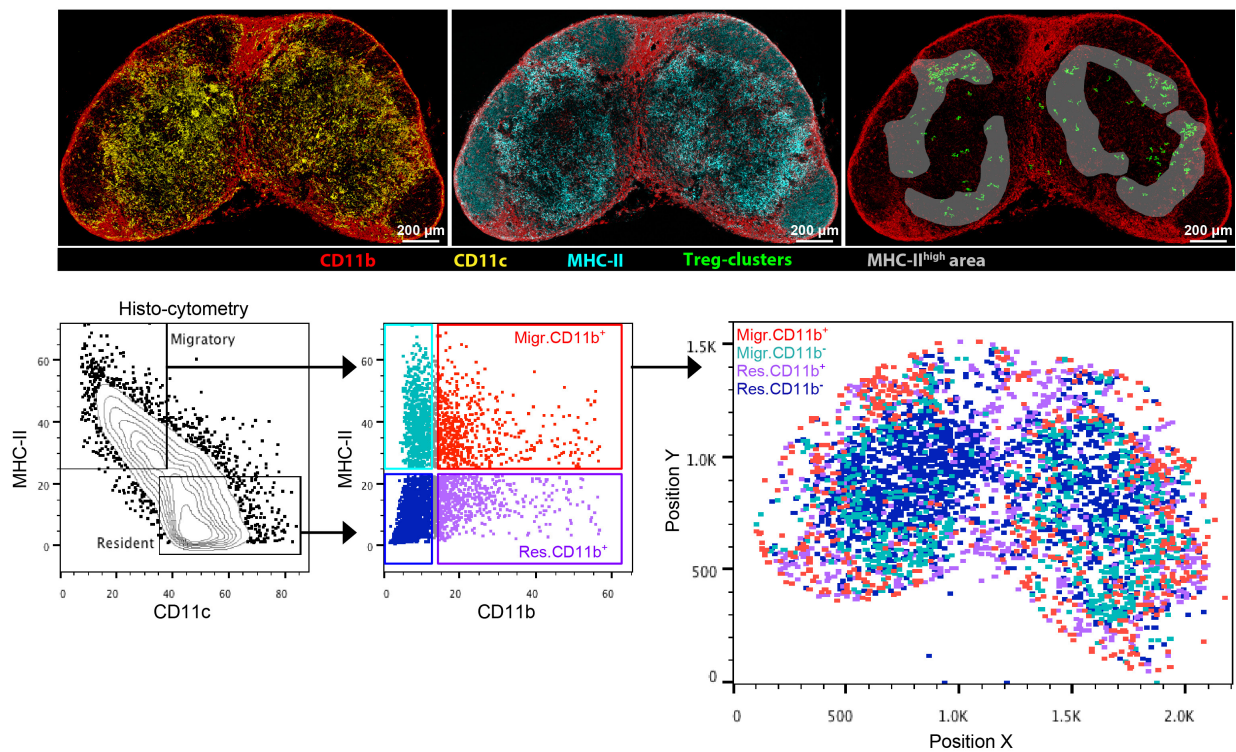


**Extended Data Figure 3 | Phenotypic characterization of IL-2-producing cells in the steady state in *Il2*<sup>WT/GFP</sup> heterozygous mice.**

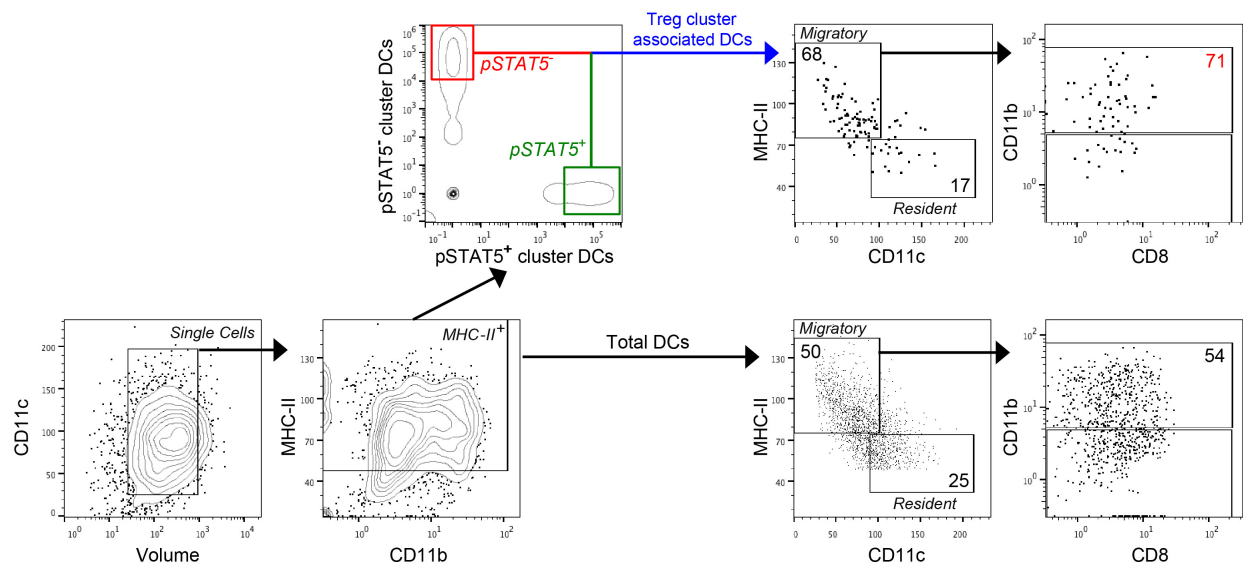
**a**, Representative flow data showing the phenotypes of cells from inguinal lymph nodes of wild-type (WT) or *Il2*<sup>WT/GFP</sup> heterozygous mice.

**b**, Quantification of the percentages of CD4<sup>+</sup> and CD8<sup>+</sup> cells within GFP<sup>+</sup>CD3<sup>+</sup> population from different lymphoid organs ( $n = 3$ , mean  $\pm$  s.e.m.). iLN, inguinal lymph node; mLN, mesenteric lymph node.

a



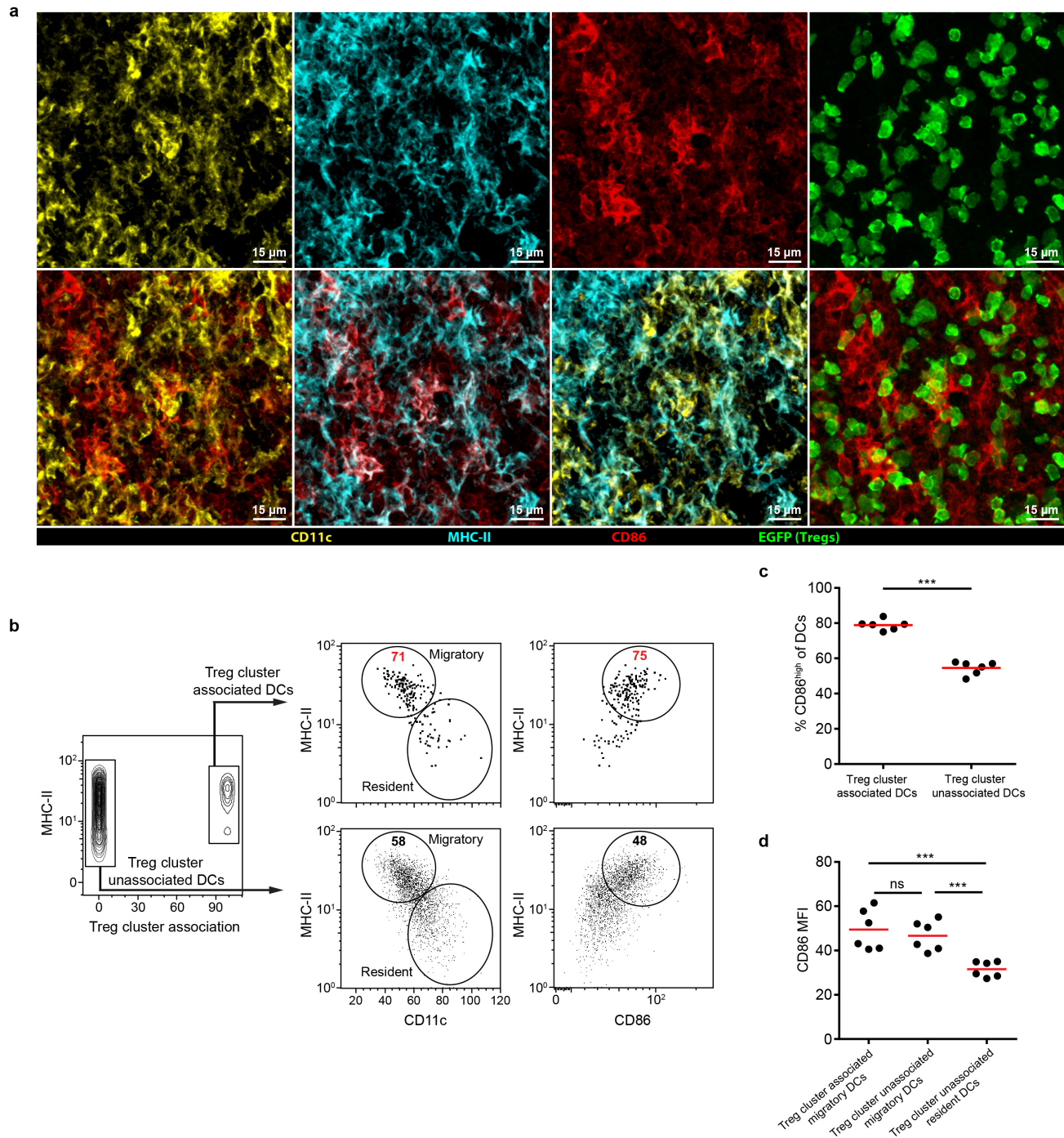
b



**Extended Data Figure 4 | Histo-cytometry analysis of dendritic cell subsets associated with  $T_{reg}$  clusters.** a, Immunofluorescence staining of an inguinal lymph node section from FOXP3-eGFP mice (upper row) and Histo-cytometry analysis of dendritic cell (DC) subsets (lower row). Surfaces for each identified dendritic cell subset are shown in Fig. 2b.

In brief, a new dendritic cell channel ( $CD11c^{+}MHC-II^{+}CD3^{-}B220^{-}$ ) was generated and used to gate on dendritic cell surface markers, as well as to create dendritic cell surfaces for subset analysis and visualization. b, Gating strategy for the phenotypic characterization of dendritic cell subsets associated with  $pSTAT5^{+}$  or  $pSTAT5^{-}$   $T_{reg}$  clusters.

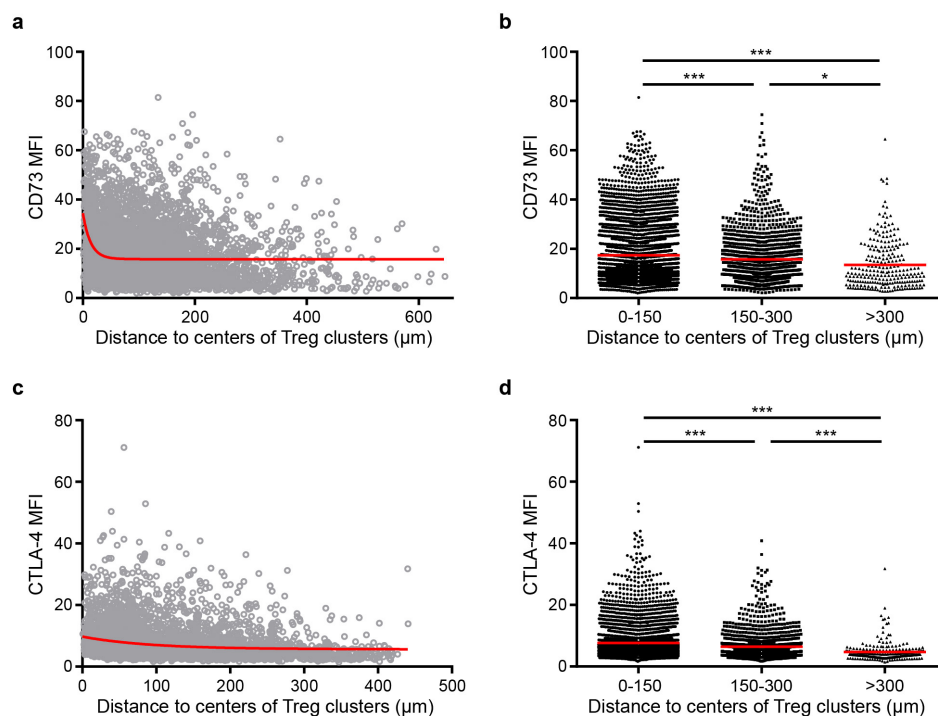




**Extended Data Figure 5 | CD86 expression on dendritic cells associated with  $T_{reg}$  clusters.** **a**, Representative images showing immunofluorescence staining of CD86 in lymph nodes from FOXP3-eGFP reporter mice.

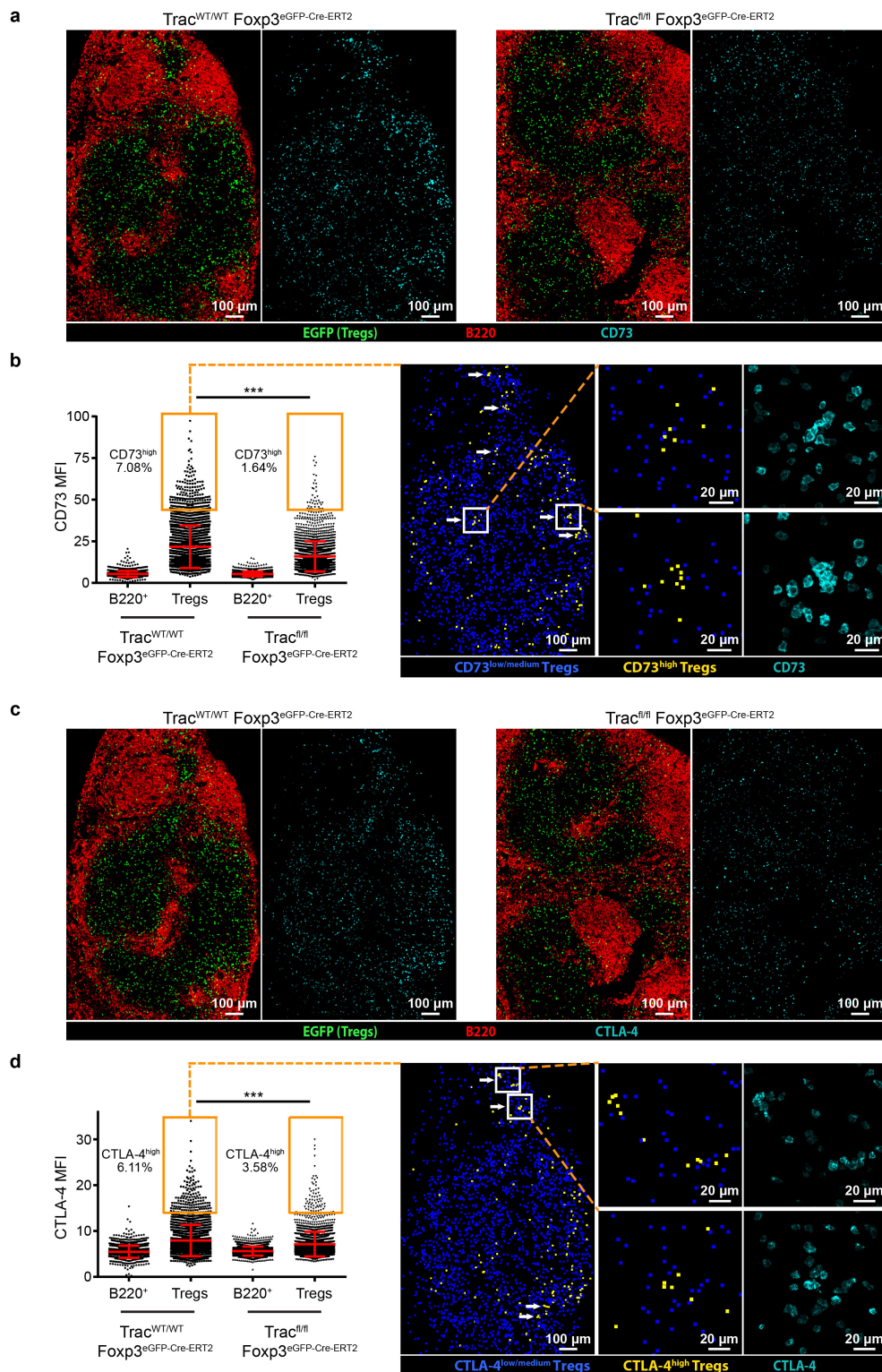
**b**, Histo-cytometry analysis of CD11c, MHC-II, and CD86 expression for dendritic cells (DCs) associated or unassociated with  $T_{reg}$  clusters.

**c, d**, Quantification of percentage of CD86<sup>high</sup> (**c**) and CD86 mean fluorescence intensity (MFI; **d**) in different dendritic cell subsets. Each dot indicates a section.  $n = 6$ , 6 sections total from 3 inguinal lymph nodes from 3 mice. \*\*\* $P < 0.001$ , as calculated by two-tailed Student's  $t$ -test (**c**) or one-way ANOVA with Tukey's post hoc test (**d**). ns, not significant.



**Extended Data Figure 6 | The correlation between CD73/CTLA4 expression on  $T_{\text{reg}}$  cells and their distance to the centre of  $T_{\text{reg}}$  clusters. a–d, Regression analysis of CD73 (a, b) and CTLA4 (c, d) based on images shown in Fig. 3e, g. MFI, mean fluorescence intensity. Each circle in**

**a and c or each dot in b and d indicates a  $T_{\text{reg}}$  cell. The red lines in a and c indicate the trend lines. The red lines in b and d indicate the mean. \*\*\* $P < 0.001$  and \* $P < 0.05$  as calculated by one-way ANOVA with Tukey's post hoc test.**



**Extended Data Figure 7 | Role of TCR signalling in high expression of CD73 and CTLA4 in T<sub>reg</sub> cells.** **a, c,** Immunofluorescence staining of CD73 (**a**) and CTLA4 (**c**) in mesenteric lymph node sections from Trac<sup>WT/WT</sup>Foxp3<sup>eGFP-cre-ERT2</sup> and Trac<sup>fl/fl</sup>Foxp3<sup>eGFP-cre-ERT2</sup> mice 7 days after tamoxifen treatment. CD73 and CTLA4 signals were gated on eGFP<sup>+</sup> T<sub>reg</sub> cells. Images are representative of lymph node sections from two mice for each group. **b, d,** Histo-cytometry analysis of CD73 and CTLA4 mean fluorescence intensity (MFI) in indicated cell populations from the

sections shown in **a** and **c**. Error bars represent mean  $\pm$  s.d. The dots in the yellow rectangles in the left panels indicate CD73<sup>high</sup> or CTLA4<sup>high</sup> T<sub>reg</sub> cells. The right panels in **b** and **d** show the distribution pattern of the gated wild-type CD73<sup>high</sup> and CTLA4<sup>high</sup> T<sub>reg</sub> cells (in left panels) in sections. Arrows indicate CD73<sup>high</sup> and CTLA4<sup>high</sup> T<sub>reg</sub> cells in representative clusters. \*\*\* $P < 0.001$  as calculated by one-way ANOVA with Tukey's post hoc test.



# Molecular structures of unbound and transcribing RNA polymerase III

Niklas A. Hoffmann<sup>1</sup>, Arjen J. Jakobi<sup>1,2</sup>, María Moreno-Morcillo<sup>1†</sup>, Sebastian Glatt<sup>1†</sup>, Jan Kosinski<sup>1</sup>, Wim J. H. Hagen<sup>1</sup>, Carsten Sachse<sup>1</sup> & Christoph W. Müller<sup>1</sup>

**Transcription of genes encoding small structured RNAs such as transfer RNAs, spliceosomal U6 small nuclear RNA and ribosomal 5S RNA is carried out by RNA polymerase III (Pol III), the largest yet structurally least characterized eukaryotic RNA polymerase. Here we present the cryo-electron microscopy structures of the *Saccharomyces cerevisiae* Pol III elongating complex at 3.9 Å resolution and the apo Pol III enzyme in two different conformations at 4.6 and 4.7 Å resolution, respectively, which allow the building of a 17-subunit atomic model of Pol III. The reconstructions reveal the precise orientation of the C82–C34–C31 heterotrimer in close proximity to the stalk. The C53–C37 heterodimer positions residues involved in transcription termination close to the non-template DNA strand. In the apo Pol III structures, the stalk adopts different orientations coupled with closed and open conformations of the clamp. Our results provide novel insights into Pol III-specific transcription and the adaptation of Pol III towards its small transcriptional targets.**

In eukaryotes, transcription of structured, small RNAs such as tRNAs, spliceosomal U6 small nuclear RNA (snRNA), ribosomal 5S RNA and 7 SL RNA is mediated by RNA polymerase III (Pol III)<sup>1</sup>. Transcriptional activity of Pol III is elevated in cancer cells, and the tumour suppressors p53 and Rb and the proto-oncogene Myc directly regulate Pol III-mediated transcription<sup>2</sup>.

Pol III is the largest of the three RNA polymerases, containing 17 subunits with a total molecular weight of 0.7 MDa, while RNA polymerase I (Pol I) and RNA polymerase II (Pol II) comprise 14 and 12 subunits, respectively. The architecture of the 10-subunit core is conserved among all eukaryotic RNA polymerases<sup>3</sup>. In Pol III it comprises the two largest subunits, C160 and C128, that form the active site. Five additional core subunits ABC27, ABC23, ABC14.5, ABC10 $\alpha$  and ABC10 $\beta$  are shared between Pol I, Pol II and Pol III. Subunits AC40 and AC19 are common between Pol I and Pol III and are closely related to their Pol II counterparts, Rpb3 and Rpb11. The Pol III core is completed by subunit C11, involved in transcription termination<sup>4,5</sup> and RNA cleavage activity<sup>5</sup>. Subunit C11 corresponds to subunit A12.2 in Pol I<sup>6</sup> and both subunits are homologous to a fusion protein that comprises the amino-terminal domain of Pol II subunit Rbp9 and the carboxy-terminal domain of TFIIS<sup>5</sup>.

The Pol III stalk comprises subunits C17 and C25 and is involved in transcription initiation<sup>3</sup> and in binding of single-stranded RNA<sup>7</sup>. The C53–C37 heterodimer is involved in transcription initiation and termination<sup>4,8,9</sup> and is distantly related to Pol I A49–A34.5 and Pol II TFIIF<sup>3,10,11</sup>. The C53–C37 heterodimer occupies a large density on the Pol III core opposite of the stalk<sup>12</sup> and extends into the DNA-binding cleft as shown by chemical crosslinking<sup>8,13</sup>. Finally, the C82–C34–C31 heterotrimer forms a stable subcomplex<sup>14,15</sup> that is positioned on the C160 clamp domain in close proximity to the stalk<sup>12,16</sup>. The C82–C34–C31 heterotrimer is involved in transcription initiation<sup>3,17</sup> and recruitment of Pol III to TFIIB<sup>18,19</sup>. Subunits C82 and C34 in the heterotrimer are distantly related to TFIIE $\alpha$  and TFIIE $\beta$ , respectively<sup>10</sup>, further supporting the hypothesis that during evolution general transcription factors such as TFIIF and TFIIE became stably associated Pol I and Pol III subunits. Maf1 is a global

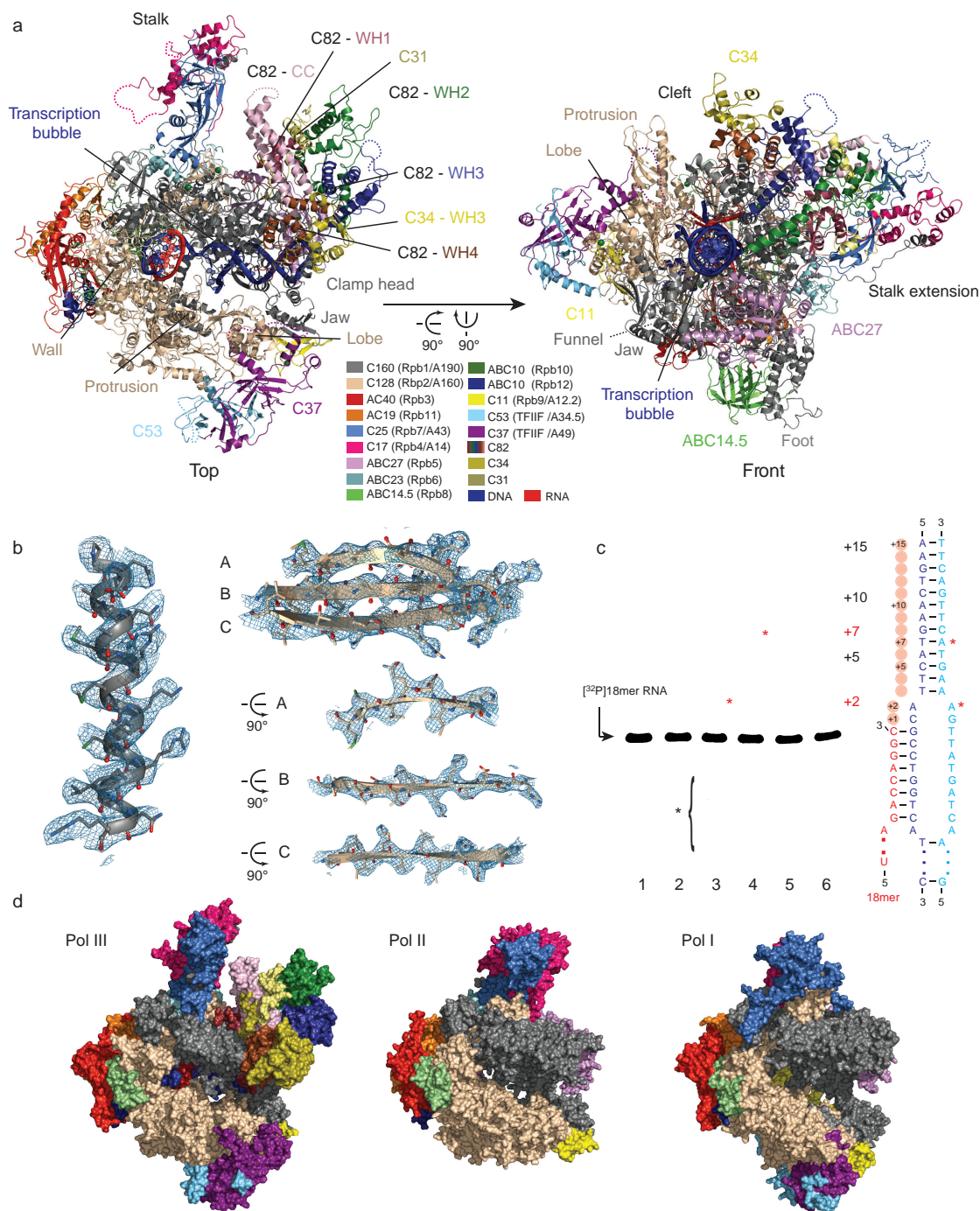
repressor of Pol III conserved in eukaryotes that under stress conditions becomes dephosphorylated and translocates to the nucleus where it directly binds Pol III<sup>20,21</sup>.

In the last decade, detailed structural studies of Pol I and Pol II contributed to elucidate basic mechanisms of DNA-dependent RNA transcription<sup>22–25</sup>. However, structural insight into Pol III-mediated transcription has been limited so far as the highest resolution Pol III structures only extend to 10 Å for apo Pol III<sup>26</sup> and to 16.5 Å and 19 Å for the Pol III elongation complex<sup>21,26</sup>. Here, we present the first atomic structures of apo Pol III and elongating Pol III obtained by single-particle electron cryo-microscopy (cryo-EM). Our reconstructions consolidate biochemical data, allow the detailed comparison with the Pol I and Pol II enzymes, and provide additional insight into Pol III-specific transcription.

## Overall architecture of Pol III

*Saccharomyces cerevisiae* Pol III was purified as previously described<sup>27</sup> yielding pure, homogenous and transcriptionally active enzyme (Fig. 1). To shed light onto the Pol III architecture and to further investigate Pol III-mediated transcription, we acquired cryo-EM images of Pol III bound to an assembled transcriptionally active bubble (elongating Pol III, Fig. 1c) and of native unbound Pol III (apo Pol III) on an FEI Titan Krios equipped with a Falcon II direct electron detector (Methods, Extended Data Fig. 1). The final map of elongating Pol III was determined using maximum-likelihood based 3D classification in RELION<sup>28</sup> at an average resolution of 3.9 Å (Extended Data Fig. 1), but extends beyond 3.5 Å in the Pol III core (Extended Data Figs 1c, 2 and 3). This electron microscopy density was used to build and refine an atomic model of the complete 17-subunit structure of elongating Pol III (Fig. 1a, Extended Data Fig. 4 and Extended Data Table 1). Representative densities (Fig. 1b, Extended Data Fig. 2) and the relatively isotropic resolution (Extended Data Fig. 1c, Extended Data Table 2) demonstrate that the quality of the cryo-EM density map is comparable to electron density maps of Pol I and Pol II obtained by X-ray crystallography at nominally higher resolutions (Extended Data Fig. 3). The overall architecture of

<sup>1</sup>European Molecular Biology Laboratory (EMBL), Structural and Computational Biology Unit, Meyerhofstrasse 1, 69117 Heidelberg, Germany. <sup>2</sup>European Molecular Biology Laboratory (EMBL), Hamburg Unit, Notkestrasse 85, 22607 Hamburg, Germany. <sup>†</sup>Present addresses: Spanish National Cancer Research Centre (CNIO), Structural Biology and Biocomputing Programme, 28029 Madrid, Spain (M.M.-M.); Malopolska Centre of Biotechnology, Jagiellonian University, Gronostajowa 7A str, 30-387 Krakow, Poland (S.G.).



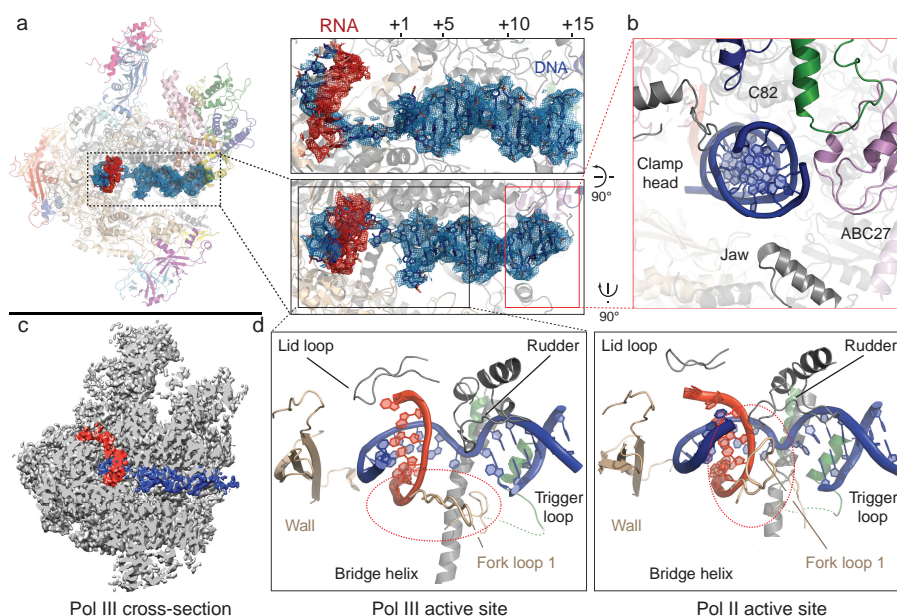
**Figure 1 | Cryo-EM structure of RNA polymerase III.** **a**, Top and front view of Pol III, individual elements and domains are labelled. Dotted lines indicate regions that are not included in the model. The colour code is presented in the corresponding boxes. **b**, Representative densities of the Pol III core with the fitted model demonstrate the high detail visible in the final cryo-EM structure. **c**, RNA extension assay demonstrates RNA elongation and cleavage activity of Pol III. The transcription bubble used for the activity assays and for the cryo-EM structure determination is depicted at the right (see also Methods). Lane 1, annealed transcription bubble with <sup>32</sup>P-labelled RNA (18mer). Lane 2, with NTP mix. Lane 3,

with Pol III but without NTPs showing the intrinsic RNA cleavage activity of Pol III; cleavage products are denoted by a black asterisk. Lanes 4 and 5, with Pol III and NTPs excluding ATP (Lane 4) or CTP (Lane 5) showing nucleotide-specificity and elongation arrest at +2 (no ATP) or +7 (no CTP) denoted by red asterisks. Lane 6, with Pol III and NTPs, the +15 run-off shows full-length extension. **d**, Surface view of the elongating Pol III structure (this study) compared to Pol I (PDB 4c3i) and Pol II (PDB 1wcm). Homologous subunits in Pol I and Pol II are coloured based on Pol III and as indicated in **a**.

the Pol III core is conserved with respect to Pol I and Pol II (Fig. 1d). However, the clamp head part of subunit C160 is enlarged compared to its Pol I and Pol II counterparts (Extended Data Fig. 5). In addition, C160 contains an extended foot that, unlike Pol I and Pol II, forms a large interface with the shared subunit ABC14.5. Furthermore, a carboxy-terminal extension of C160, unique to Pol III, protrudes

from the core and together with the C160 N terminus contacts the stalk. Similarly, the second largest subunit C128 shows an overall conserved fold, but contains an extended protrusion that increases the depth of the DNA-binding cleft in comparison to Pol I and Pol II (Extended Data Fig. 5). The Pol III cryo-EM structure also includes the Pol III-specific C82–C34–C31 heterotrimer and the C53–C37





**Figure 2 | Transcription of Pol III and association with DNA–RNA duplex.** **a**, The left panel shows an elongating Pol III ribbon model with the segmented density of the transcription bubble displayed at 4.5 Å for better visibility. DNA and RNA densities are shown in blue and red, respectively. The transcription bubble is shown in stick representation. The downstream DNA duplex is embedded in the cleft, an eight-base-pair DNA–RNA hybrid was built based on the density. The two right panels show close-ups at two orthogonal views. **b**, The downstream DNA duplex is tightly bound between the jaw (grey), the clamp head (grey), subunit ABC27 (pink) and two C82 WH domains (WH2, green and WH3, blue). **c**, Cross-section of the elongating Pol III density at 3.9 Å with

coloured density corresponding to the transcription bubble. The density for the downstream DNA duplex exceeds even beyond bp +15 in the map, although it is much weaker compared to DNA density in the cleft, thus likely corresponding to an unstably stacked second DNA duplex. **d**, Close-up view of the active site of Pol III (left) and Pol II (right, PDB 1y1w). An extended rudder that points towards a stretch of the protrusion (residues 390–400) and a buried fork loop 1 suggest that the DNA–RNA hybrid in the Pol III core is less tightly bound compared to Pol II (right panel), where fork loop 1 protrudes into the core and together with the rudder and the wall forms a barrier. Similar as in Pol II, the trigger loop is unstructured in Pol III.

heterodimer, both showing several unexpected features as discussed later.

The apo Pol III data set yielded two major 3D classes showing distinct conformations at 4.6 Å and 4.7 Å resolution (Extended Data Fig. 1b, c). One reconstruction is very similar to the elongating Pol III (root mean square deviation (r.m.s.d.) = 0.43 Å<sub>4,813 C $\alpha$  atoms aligned</sub>), whereas the second reconstruction shows an altered orientation of the stalk, the heterotrimer and a more open cleft resulting in a larger difference with elongating Pol III (r.m.s.d. = 2.73 Å<sub>4,795 C $\alpha$  atoms aligned</sub>). The two apo Pol III conformations presumably result from ‘closed’ and ‘open’ states of the clamp domain as discussed below (Extended Data Fig. 6a–c).

### A narrow cleft encloses DNA

In both apo Pol III and the elongating Pol III structures we observe a characteristic narrower cleft in comparison to Pol I and Pol II (Extended Data Fig. 6d). The DNA duplex is embedded into the Pol III cleft and extends from downstream base pair (bp) +14 towards the active site until the upstream bp –9 (Fig. 2a). The DNA duplex is anchored between the jaw and lobe domains and the oppositely positioned extended clamp head. In addition, the WH2 and WH3 domains of C82 (see below) lie in close proximity and further stabilize downstream DNA (Fig. 2b). Subunit ABC27 completes this enclosure by inserting a proline-containing loop into the minor groove between bps +11 to +14, thus threading the DNA duplex towards the active centre. In the crystal structure of elongating Pol II bound to a transcription bubble that comprises upstream and downstream DNA duplexes (53-mer DNA oligonucleotides)<sup>29</sup>, the proline-containing loop also protrudes into the minor groove of downstream DNA, while upstream DNA interacts with a hairpin-loop/wedge in Pol II subunit Rbp2 (residues 862–874) also present in Pol III subunit C128 (residues 794–806) as well as in Pol I. Global recognition of upstream and downstream DNA

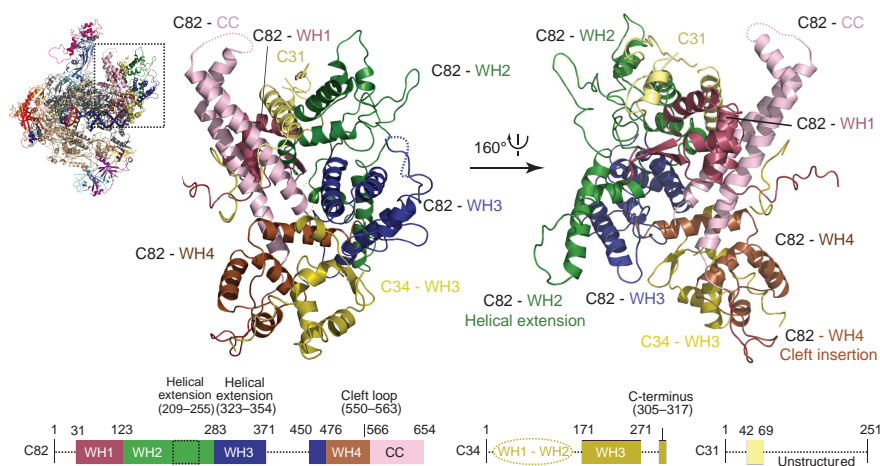
therefore appears to be conserved among the three eukaryotic RNA polymerases.

The template DNA strand unwinds at bp +2 and a characteristic A-type DNA–RNA hybrid forms at positions –1 to –9 (Fig. 2a). Additional density in the RNA exit channel likely corresponds to the emerging and thus more flexible single-stranded RNA (Fig. 2c). No clear density for upstream DNA is visible, indicating flexibility of the emerging duplex. Another noticeable feature is the strong density of the downstream DNA double strand in comparison to the much weaker density for the DNA–RNA duplex (Fig. 2a–c). This is in contrast to the Pol II elongation complex, where both densities are of equal quality<sup>29–31</sup>. Remarkably, the DNA–RNA duplex is tightly associated with the Pol II elongation complex between the wall and fork loop 1/rudder, whereas in Pol III the rudder and fork loop 1 reach towards the protrusion, thereby reducing the association of the DNA–RNA hybrid with the Pol III active centre (Fig. 2d). While the DNA–RNA hybrid appears to be only loosely associated with the active site, we observe a tight enclosure of the downstream DNA duplex at the entrance of the DNA-binding cleft.

### Pol III heterotrimer protrudes into the cleft

Although the crystal structure of the human C82 orthologue hRPC62 (ref. 32) is available, its integration into the C82–C34–C31 heterotrimer and the precise orientation of the heterotrimer within the complete Pol III enzyme could not be clarified by previous cryo-EM reconstructions owing to their limited resolution<sup>12,21,26</sup>. In addition, the functional roles of the seven winged-helix (WH) domains present in subunits C82 and C34 and often found in transcription factors as DNA-binding or protein–protein interaction modules<sup>33</sup> are still poorly understood. Our cryo-EM structure shows how the heterotrimer packs onto the clamp head by forming a large, hydrophobic interface through various WH domains (C82–WH1/WH4 and C34–WH3; Fig. 3). In this case the WH domains serve as protein–protein interacting domains, although they





**Figure 3 | Architecture of the Pol III-specific heterotrimer.** The panel shows the C82–C34–C31 heterotrimer as ribbon representation in top (left) and bottom (right) view. Schematic representations of C82 (left), C34 (middle) and C31 (right) depict domain boundaries. Structured and disordered regions are marked with solid and dotted lines, respectively.

might still contact DNA during open complex formation in the transcription initiation process. The coiled-coil domain of C82 protrudes towards the stalk and the remaining C82-WH2 and WH3 are facing away from the core. Notably, C82-WH2 and C82-WH3 align with the clamp head to reach around bp +15 towards downstream DNA. Furthermore, an additional long ‘cleft loop’ extending from C82-WH4 passes through a canyon in the clamp head into the DNA-binding cleft close to DNA bp +7 (Extended Data Fig. 7a). Subunit C34 comprises three WH domains that span from C82 towards the protrusion crossing the DNA-binding cleft<sup>16</sup>. TFIIE $\alpha/\beta$  also possesses three WH domains and, in the cryo-EM structure of the human Pol II pre-initiation complex, crosses the DNA-binding cleft<sup>34</sup>. The two N-terminal C34 WH domains are not visible in the Pol III cryo-EM density like the A49 tandem WH domain in the Pol I crystal structure<sup>24,25</sup>, while the third WH domain tightly associates with C82 and is located at the periphery of the heterotrimer. The C34 C-terminal region following the third WH domain passes the C82-WH4 and contacts the C82 coiled-coil domain (Fig. 3). The third subunit of the heterotrimer, namely C31, was predicted to be largely unstructured and associated between C82 and the stalk, as shown by crosslinking analysis<sup>16</sup>. However, we were able to build a mainly helical element of C31 (residues 42–69), demonstrating that C31 extends along the surface from the C34 C terminus towards C82-WH4 over the C82 coiled-coil domain to reach the stalk, where it becomes disordered. Additional density close to the interface of C82-WH1 and WH2 and between the heterotrimer and the stalk is most likely to correspond to C-terminal stretches of C31, but no sequence could be unambiguously assigned to it (Extended Data Fig. 7b). Nevertheless, the topology of C31 and the extended interaction interface with at least three WH domains, the C82 coiled-coil domain and the stalk confirm the previously reported role of C31 in connecting the heterotrimer to the Pol III core and stalk<sup>17,35,36</sup>.

### Transcription termination by Pol III heterodimer

Previous electron microscopy and crosslinking studies of Pol III and the C53–C37 subcomplex positioned a conserved dimerization module at the lobe of Pol III<sup>12,13,21,26</sup> similar to TFIIF and A49–A34.5 in Pol II<sup>37</sup> and Pol I<sup>24,25</sup>, respectively. On the other hand, extensions of C53 and C37 crosslink close to the active site, the stalk and the heterotrimer<sup>8,13,16</sup>. Our structure better characterizes the interaction network of the C53–C37 heterodimer with other Pol III subunits. Notably, C37 shows an extended contact surface with C11 and the cleft when compared to TFIIF $\alpha$  and A49 (Fig. 4a), consistent with its proposed role in C11 association with the core<sup>9</sup>.

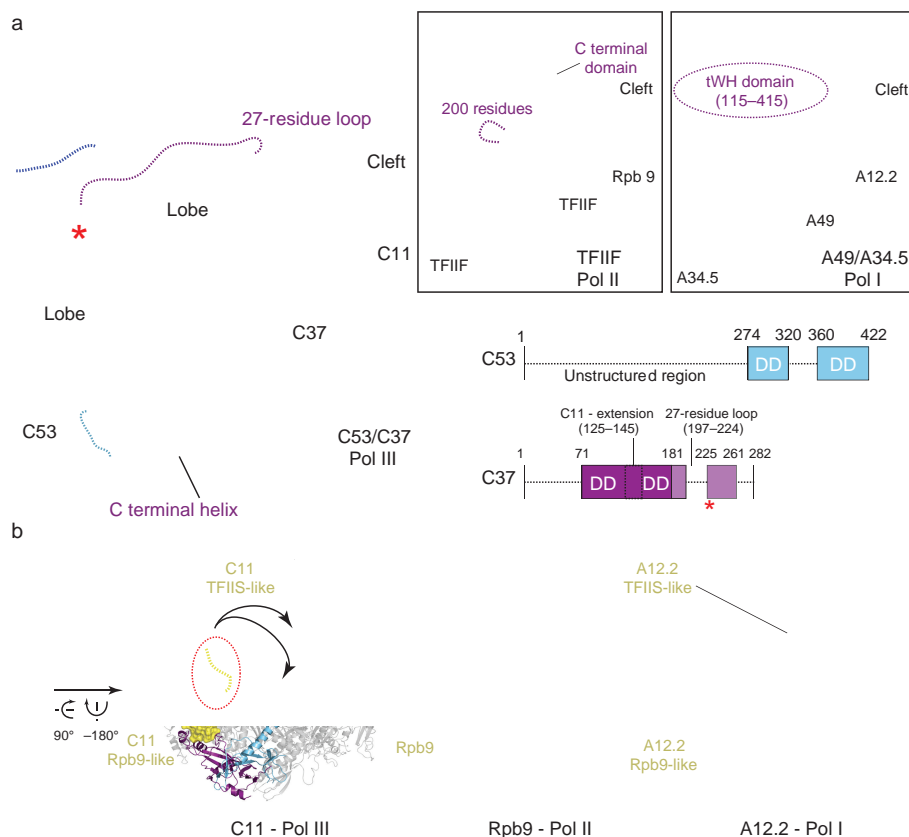
Our cryo-EM structure also rationalizes the role of the C53–C37 heterodimer in Pol III transcription termination that only requires a stretch of 5–7 thymines in the non-template DNA strand for efficient transcription termination<sup>38</sup>. Subunit C37 extends towards the DNA-binding cleft where it positions a flexible loop (residues 197–224) that

has been shown to contact C34 and the Pol III-specific TFIIB subunit Bdp1<sup>13,39</sup> before folding back into a helix (residues 230–240; Extended Data Fig. 8a). Deleting the five residues (R226, L227, T228, G229, S230) leading into this helix produces a terminator read-through phenotype in *S. cerevisiae*<sup>4</sup>, and in *Schizosaccharomyces pombe* the corresponding region was identified as a hotspot for terminator read-through mutations<sup>40</sup>. The same residues have also been cross-linked to C128 (ref. 13) and in the cryo-EM structure are packed onto a section of the C128 lobe that when deleted also results in a termination read-through phenotype<sup>41</sup>. In addition, these five residues are in close proximity to the non-template DNA strand that is flexible in the Pol III structure (Extended Data Fig. 8a). Direct interactions between subunit C37 presumably involve these five residues and the first four thymines of the non-template DNA strand and result in a conformational switch of Pol III towards a metastable pre-termination complex, while the presence of a fifth thymine results in transcription termination<sup>4</sup>. In contrast, Pol I and Pol II require additional *cis*- and/or *trans*-acting factors for transcription termination<sup>42</sup>, and accordingly TFIIF or A49–A34.5 do not possess similar features (Fig. 4a).

The C terminus of C37 folds back and forms a hydrophobic interface with the C53 and C37 core consistent with the crucial role of the C37 C terminus in assembling and stabilizing the C53–C37 heterodimer<sup>9</sup>. C53 is located below the lobe and extension 2 and its C-terminal dimerization domain is indeed tightly associated with the C37 dimerization domain (Fig. 4a). The remaining 270 residues of the N-terminal region of C53 are flexible and could not be attributed to any density, although crosslinking experiments suggest that they are involved in a complex interaction network<sup>8,13</sup>. Interestingly, weak (but continuous) density in the cleft between C11 and C37 could be identified as a part of C53 based on specific photo-crosslinks<sup>13</sup>, although the density is too weak for model building and sequence assignment (Extended Data Fig. 8b). This observation implies that the N-terminal extension of C53 travels from the C160 funnel back to the periphery of the heterodimer in close proximity to the previously described C37–C11 interacting region before moving towards the cleft, the stalk and the heterotrimer<sup>8,13,16</sup>.

### The TFIIS-like domain of C11 is flexible

The core subunit C11 is responsible for the intrinsic RNA cleavage activity of Pol III<sup>43</sup> that is likely to be mediated by its conserved TFIIS-like domain<sup>5</sup>. Notably, the orthologous Pol I subunit A12.2 also contains a C-terminal TFIIS-like domain that associates to the active site<sup>24,25</sup>. Both C11 and A12.2 also harbour a conserved N-terminal Zn-binding domain with high homology to Pol II Rbp9. The N-terminal Rbp9-like domain of C11 is anchored between C37, the C128 lobe domain and the jaw domain of subunit C160 at a similar position as the N-terminal domains of Rbp9 and A12.2 (Fig. 4b). In contrast, density corresponding to the C-terminal TFIIS-like domain was not observed in the elongating Pol III structure, but was only identified in both apo Pol III structures



**Figure 4 | Architecture and Pol III-specific function of the C53–C37 heterodimer and C11.**  
**a**, Model of the Pol III C53–C37 heterodimer shown in ribbon representation bound to the Pol III core (left), Pol II homologue TFIIFα/β bound to the Pol II core (centre, PDB 4v1n) and Pol I homologue A49–A34.5 bound to the Pol I core (right, PDB 4c3i). The red asterisk (left panel) marks the position of the five residues that upon deletion lead to a terminator read-through phenotype<sup>13</sup>. Schematic representations of C53 and C37 show the domain boundaries of the dimerization domain (DD) and additional elements. Dotted lines indicate unstructured regions. **b**, Conformation of subunit C11 in Pol III (left), subunit Rpb9 in Pol II (middle), and subunit A12.2 in Pol I. Subunits C11, Rpb9 and A12.2 are depicted with yellow surface rendering; C53–C37, TFIIFα/β and A49–A34.5 are depicted in ribbon representation, all other subunits are coloured in grey. Arrows indicate the potential movement of the C11 C-terminal TFIIS domain, the red dotted circle indicates the linker that connects the C11 N- and C-terminal domains.

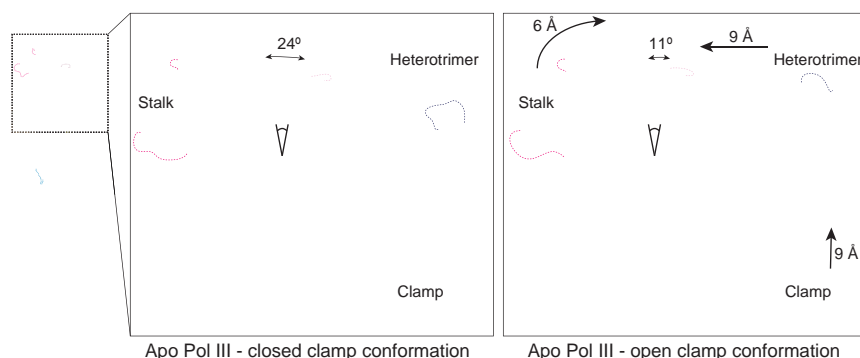
where it occupies a cleft between Rpb5 and the C160 funnel as observed at a low density threshold (Extended Data Fig. 7c). The position of the C11 C-terminal domain is far away from the position of the corresponding A12.2 C-terminal domain, suggesting that the C11 C-terminal domain is mobile and only temporally recruited to the catalytic centre (in analogy to TFIIS in Pol II). A long linker (residues 37–61) that connects both domains presumably accounts for the required mobility (Fig. 4b). C11 mutations at the extended interface between C37 and C11 N-terminal domain induce terminator read-through transcripts, while mutations of the C-terminal domain of C11 affect RNA 3' cleavage, reflecting the involvement of the flexibly linked N- and C-terminal domains of C11 in two termination-related, yet distinct activities<sup>44</sup>.

### Pol III stalk relays conformational changes

The Pol III stalk subunits C25 and C17 are homologous to Pol II Rpb7–Rpb4 and Pol I A43–A14 (refs 7, 36). Consistent with previous low-resolution electron microscopy studies<sup>12,21,26</sup>, the Pol III cryo-EM structure confirms that the closed HDRC domain fold observed in the *S. pombe* C25–C17 heterodimer<sup>45</sup> (PDB 3ayh) describes the conformation of the stalk on active Pol III most appropriately. Furthermore, we see a tight interaction network of the Pol III stalk and the C160 N- and C-terminal extensions, the latter contacting and positioning the C17 HDRC domain

on C25 (Extended Data Fig. 8c). This leads to a tight anchoring of the stalk to the core, which is additionally strengthened by the Pol III-specific helix of C25 that extends from the stalk and contacts the clamp.

Remarkably, the open and closed clamp conformations of apo Pol III (Extended Data Fig. 6b) demonstrate two structurally distinct conformations of the stalk, the clamp head and the heterotrimer (Fig. 5). A flexible clamp has been also reported in the bacterial RNA polymerase, where it was shown that the clamp is predominantly open in the unbound conformation, then closes during initiation and elongation<sup>46</sup>. Furthermore, the archaeal RNA polymerase and Pol II contain a flexible clamp and in both systems the status of the clamp is associated with the stalk<sup>30,34,47</sup>. In contrast to archaeal RNA polymerase and Pol II, in Pol III the clamp movement is less pronounced, resulting in a narrower Pol III cleft compared to other RNA polymerases even in the open clamp conformation (Extended Data Fig. 6c, d). Here, we show both clamp conformations in Pol III visualized from the same sample and speculate that a moving stalk in Pol III can mediate the observed conformational changes. The open clamp conformation opens the cleft and Pol III could thus better associate with target DNA, whereas a transition to the closed clamp positions the heterotrimer next to the unwound DNA and could enable C82 and C34-mediated promoter opening and subsequent elongation.



**Figure 5 | Conformational changes in apo Pol III.** 'Closed clamp' (left) and 'open clamp' (right) conformations of apo Pol III. The C82–C34–C31 heterotrimer, the stalk and the clamp are shown in ribbon representation, the core in surface representation. Arrows in the right panel and corresponding values indicate movements of the stalk, the heterotrimer and the clamp head relative to the closed clamp state.

# Sublimation in bright spots on (1) Ceres

A. Nathues<sup>1</sup>, M. Hoffmann<sup>1</sup>, M. Schaefer<sup>1</sup>, L. Le Corre<sup>1,2</sup>, V. Reddy<sup>1,2</sup>, T. Platz<sup>1</sup>, E. A. Cloutis<sup>3</sup>, U. Christensen<sup>1</sup>, T. Kneissl<sup>4</sup>, J.-Y. Li<sup>2</sup>, K. Mengel<sup>5</sup>, N. Schmedemann<sup>4</sup>, T. Schaefer<sup>1</sup>, C. T. Russell<sup>6</sup>, D. M. Applin<sup>3</sup>, D. L. Buczkowski<sup>7</sup>, M. R. M. Izawa<sup>3,8</sup>, H. U. Keller<sup>9</sup>, D. P. O'Brien<sup>2</sup>, C. M. Pieters<sup>10</sup>, C. A. Raymond<sup>11</sup>, J. Ripken<sup>1</sup>, P. M. Schenk<sup>12</sup>, B. E. Schmidt<sup>13</sup>, H. Sierks<sup>1</sup>, M. V. Sykes<sup>2</sup>, G. S. Thangjam<sup>1</sup> & J.-B. Vincent<sup>1</sup>

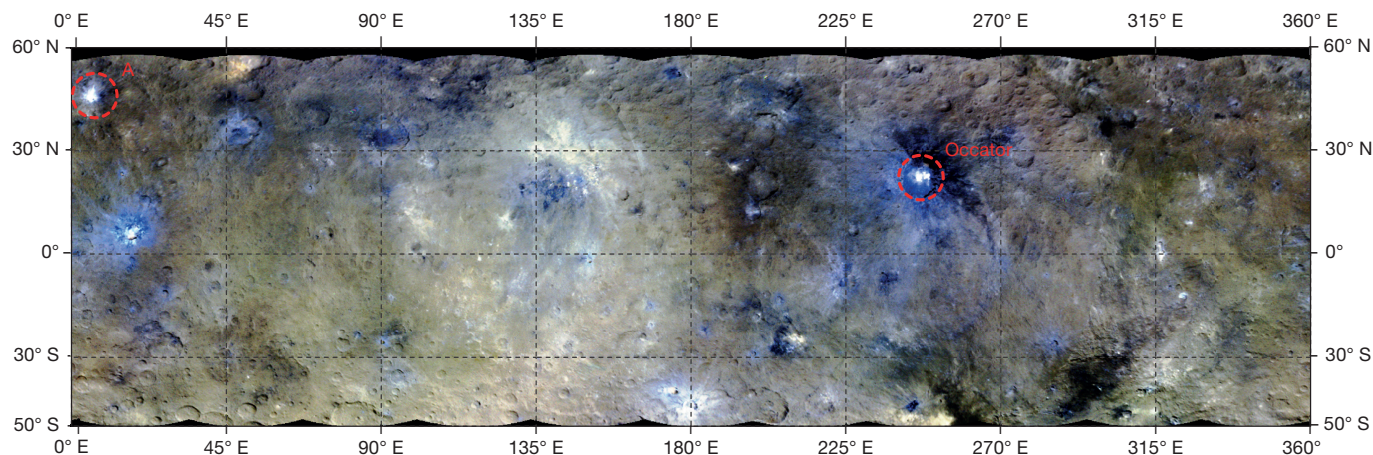
The dwarf planet (1) Ceres, the largest object in the main asteroid belt<sup>1</sup> with a mean diameter of about 950 kilometres, is located at a mean distance from the Sun of about 2.8 astronomical units (one astronomical unit is the Earth–Sun distance). Thermal evolution models suggest that it is a differentiated body with potential geological activity<sup>2,3</sup>. Unlike on the icy satellites of Jupiter and Saturn, where tidal forces are responsible for spewing briny water into space, no tidal forces are acting on Ceres. In the absence of such forces, most objects in the main asteroid belt are expected to be geologically inert. The recent discovery<sup>4</sup> of water vapour absorption near Ceres and previous detection of bound water and OH near and on Ceres (refs 5–7) have raised interest in the possible presence of surface ice. Here we report the presence of localized bright areas on Ceres from an orbiting imager<sup>8</sup>. These unusual areas are consistent with hydrated magnesium sulfates mixed with dark background material, although other compositions are possible. Of particular interest is a bright pit on the floor of crater Occator that exhibits probable sublimation of water ice, producing haze clouds inside the crater that appear and disappear with a diurnal rhythm. Slow-moving condensed-ice or dust particles<sup>9,10</sup> may explain this haze. We conclude that Ceres must have accreted material from beyond the ‘snow line’<sup>11</sup>, which is the distance from the Sun at which water molecules condense.

We observed Ceres using the Framing Camera (FC) on board the Dawn spacecraft. The FC is equipped with a panchromatic (clear) filter and seven colour filters, covering the wavelength range 0.4  $\mu\text{m}$  to 1.0  $\mu\text{m}$  (ref. 8). Image calibration and spectral cube computation were performed as for the Vesta phase of the Dawn mission<sup>12,13</sup>.

We find that the general surface of Ceres is rather dark (similar in brightness to fresh asphalt) and spotted with bright patches (ranging between the brightness of concrete and that of ocean ice). Figure 1 shows several bright spots across the surface. The inner part of the central spot on the floor of the Occator crater is the brightest surface on Ceres, with an absolute reflectance (converted from image reflectivity data by Hapke modelling) of  $\sim 0.25$ ; the second-brightest feature (absolute reflectance  $\sim 0.07$ ) is tentatively named feature A, and is associated with a 10-km-diameter crater far from Occator. We counted more than 130 bright spots on Ceres, most of them associated with impact craters (Extended Data Fig. 1).

The floor of Occator (diameter 90.5 km, depth 4 km) exhibits multiple, partially coalesced bright spots (Fig. 2a, Extended Data Figs 2 and 4). The largest of these spots corresponds to a central pit ( $\sim 10$  km wide,  $\sim 0.5$  km deep) covered by bright material. The available data do not allow a conclusion to be reached on the local depths of this layer. The pit is traversed by dark lineaments which are possibly fractures, and seems to be surrounded by remnants of a central peak up to 0.5 km high. The morphology of Occator suggests that it is a relatively young crater, as the crater rim and walls appear rather sharp, and there are abundant terraces and landslide deposits. Very few superposed impact craters are observed on its rim, floor and ejecta. Occator partially formed on an east-southeast-trending fault, causing the rim scallop. Crater-based dating yields a relatively young crater age of  $\sim 78$  Ma (Extended Data Fig. 3).

Feature A is located in the northwest part of an ancient 143-km-diameter crater (Fig. 2b, Extended Data Fig. 5) overlapped by a degraded 40.5-km-diameter crater. More specifically, the bright

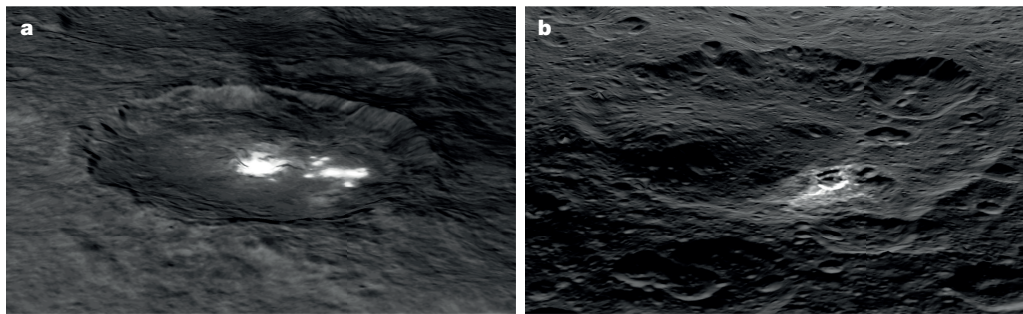


**Figure 1 | Enhanced colour mosaic of the surface of Ceres.** Data were obtained from 14,000 km distance and projected equirectangular in an RGB mosaic (colour coding of wavelengths: red R = 0.96  $\mu\text{m}$ , green

G = 0.75  $\mu\text{m}$ , blue B = 0.44  $\mu\text{m}$ ). Red circles mark the location of the two brightest spots (Occator and feature A). Several dark patches, primarily associated with crater ejecta blankets, are also visible.

<sup>1</sup>Max Planck Institute for Solar System Research, Goettingen, Germany. <sup>2</sup>Planetary Science Institute, Tucson, Arizona, USA. <sup>3</sup>University of Winnipeg, Winnipeg, Canada. <sup>4</sup>Freie Universitaet Berlin, Berlin, Germany. <sup>5</sup>Technische Universitaet Clausthal, Clausthal-Zellerfeld, Germany. <sup>6</sup>University of California, Los Angeles (UCLA), Los Angeles, California, USA. <sup>7</sup>Johns Hopkins University, Laurel, Maryland, USA. <sup>8</sup>Royal Ontario Museum, Toronto, Canada. <sup>9</sup>TU Braunschweig, Braunschweig, Germany. <sup>10</sup>Brown University, Providence, Rhode Island, USA. <sup>11</sup>Jet Propulsion Laboratory, Pasadena, California, USA. <sup>12</sup>Lunar and Planetary Institute, Houston, Texas, USA. <sup>13</sup>Georgia Institute of Technology, Atlanta, Georgia, USA.





**Figure 2 | Perspective views of brightest spots on Ceres.** **a**, These spots are located on the floor of Occator crater at 19.4° N, 239.0° E (**a**) and at an unnamed crater at 41.7° N, 0.8° E (feature A; **b**). The clear filter images are wrapped onto topographic information derived from a stereo pair of images with elevation exaggerated three times (resolution  $\sim 0.4$  km per

pixel). The brightest spot on Ceres (**a**) is located in a depression on the floor of the  $\sim 90.5$ -km-diameter Occator crater. More than 10 smaller and fainter spots are found east of the central spot. The second brightest region on Ceres (**b**) is located in and around a 10-km-diameter crater (feature A).

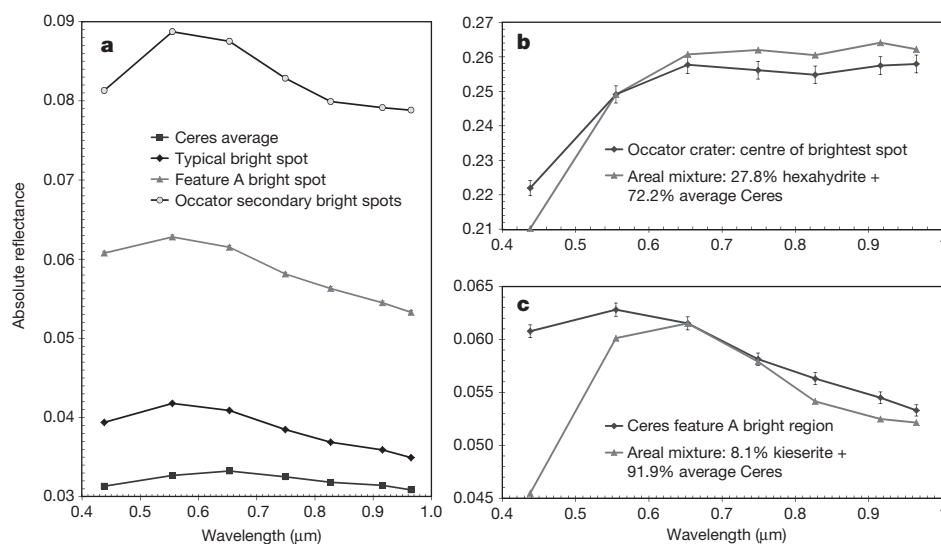
material occurs inside a young, 10-km-diameter crater and nearby. Some spots located between 290° E and 360° E appear to be related to tectonic lineaments and/or large domes.

Because of the absence of strongly diagnostic absorption features in the wavelength range of the FC, any identification of specific phases must be considered tentative. We used absolute reflectance and the shape of the spectra to constrain the composition of the bright spots. These spots show not only substantially higher reflectances, but also a spectral shape that is different from the average for the whole of Ceres (see Fig. 3a, Extended Data Fig. 6). The spots exhibit a wavelength of maximum reflectance that is shifted from  $0.65\ \mu\text{m}$  (average surface) to  $0.55\ \mu\text{m}$ ; also, the reflectance peak is more pronounced for the bright spots. It is important to note that the central part of the brightest Occator spot exhibits a spectrum that is different from the other bright spot spectra (Fig. 3b). Previous ground-based spectral observations of Ceres have suggested its surface to be similar to carbonaceous chondrite meteorites<sup>14</sup>. Observations by Dawn's Visible and Infrared Spectrometer have now revealed an average global surface that contains ammoniated phyllosilicates<sup>15</sup>. The absolute reflectance of the small-scale Occator bright spots coupled with the Herschel observatory's water vapour detection<sup>4</sup> suggests three possible candidate materials: water ice; iron-depleted clay minerals; and salts.

We conducted an extensive analysis of potential analogue materials. Of all materials considered, the centre of the brightest spot in Occator

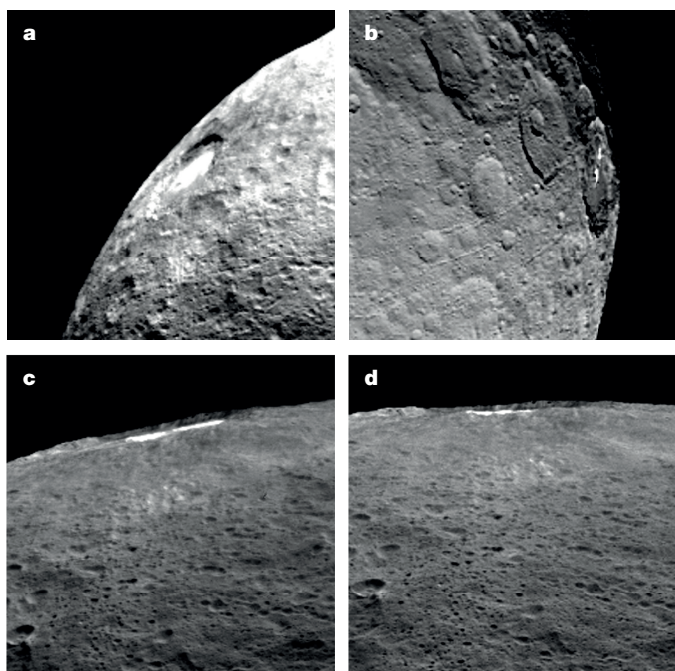
matches best with hexahydrite (six-hydrated magnesium sulfate, Fig. 3b). It is less consistent with other plausible minerals like smectites, although additional components that are spectrally featureless in this wavelength range, such as water ice, may also be present. Because of the detection of haze in the vicinity of the brightest spots (see below), a volatile component must be involved, which is likely to be water, as Herschel observations suggest<sup>4</sup>. With increasing distance from the centre of the Occator pit, the spectra become similar to those of the less bright spots on the surface, shifting the wavelength of maximum reflectivity from  $0.65\ \mu\text{m}$  to  $0.55\ \mu\text{m}$ , and increasing the negative spectral slope on the long-wavelength side of  $0.55\ \mu\text{m}$ . This spectral behaviour can be explained by a decreasing contribution from the brightest materials, and a possible change in their composition—to a less hydrated magnesium sulfate (kieserite, a mono-hydrated magnesium sulfate; Fig. 3c). Because of the lack of diagnostic spectral features in many candidate minerals, we stress that we cannot rule out other materials or processes, such as localized heating in contact with an ice-rich sub-surface material that leads also to the formation of an iron-poor clay mineral, similar to processes on Mars<sup>16,17</sup>.

In April 2015 the FC searched for plume activity; during this search, Ceres' limb was observed at high phase angles up to  $155^\circ$ . The goal was to detect potential dust particles in forward light scattering geometry. Although we did not detect any obvious plumes emanating from Ceres' lit limb, we observed near-surface haze in Occator that showed diurnal



**Figure 3 | Ceres colour spectra.** **a**, Bright spot spectra peak around  $0.55\ \mu\text{m}$  and show negative spectral slopes on the long-wavelength side of  $0.55\ \mu\text{m}$ , whereas the average Ceres spectrum is darker and peaks around  $0.65\ \mu\text{m}$ . **b**, Spectrum of the centre of the brightest Occator spot compared with a numerical combination of an average Ceres spectrum plus a

hexahydrite laboratory spectrum (see Methods section 'Spectral analysis'). **c**, Spectrum of feature A compared with a combination of an average Ceres spectrum and kieserite. The difference between the spectrum of feature A and this mixture at  $0.44\ \mu\text{m}$  is probably due to a difference in iron-bearing mineral content. Error bars,  $\pm 1$  s.d.



**Figure 4 | Views of Occator crater at different times.** Occator contains the brightest spot on Ceres. This is a pit covered by bright material, and the surrounding area shows a specific diurnal brightness rhythm, which becomes detectable at oblique views. **a**, Oblique view at noon reveals a diffuse near-surface haze (white) that fills the floor of its host crater. This haze disappears completely at dusk (**b**). The low column density of the haze is indicated by the very oblique limb views at noon (**c** and **d**), and the haze does not extend above the elevated southwestern part of the crater floor (left in **c**). There is a lapse of 450 s between the taking of the images shown in **c** and **d**.

variation, being pronounced at local noon and disappearing completely near dusk (see Fig. 4, Extended Data Fig. 7). Common among all four images displayed in Fig. 4 is an emission angle (angle between surface normal and line of sight) of  $71^{\circ}$ – $87^{\circ}$ , causing a high optical depth<sup>18</sup>. Despite the same range of optical depth at dusk, the absence of a detectable brightness excess at that time (Fig. 4b) proves the haze hypothesis. Any brighter solid material on the surface would remain at least partially visible at its phase angle of only  $36^{\circ}$ , which is even lower than that for the other three images ( $43^{\circ}$ – $70^{\circ}$ ).

Given the geographical correlation between Occator and a localized water vapour source detected by Herschel<sup>4</sup>, we suggest that this haze layer consists of water-ice particles and dust. Similarly to the activity at the surface of a comet, the haze forms in the morning after sunrise by the sublimation of water. The stream of vapour lifts tiny particles of dust and residual ice. Despite Ceres' lack of a global internal heat source, this process resembles in part a mechanism suggested for vents on Enceladus<sup>10</sup>. We conclude that a similar process occurs at feature A (see Fig. 2b). The diurnal variation of haze on Ceres could also be responsible for the short time variability of the water vapour absorption measured by the Herschel telescope<sup>4</sup>. The presence of near-surface haze is realistic, because the measured outflow velocity of water on Ceres<sup>4</sup> is  $0.3$ – $0.7 \text{ km s}^{-1}$ , comparable to the escape velocity<sup>4</sup> of  $\sim 0.52 \text{ km s}^{-1}$ , which matches the expected speed for water molecules in this environment. As water sublimates it expands away from the surface, dragging along dust. Because of the speed distribution in the gas, a significant fraction will not leave Ceres, causing short-term haze and possibly proximal deposition. By assuming a system without internal forces and considering a water vapour loss of a few kilograms per second during daytime<sup>4</sup>, the Occator spots are among the youngest features on Ceres. Our non-detection of plume activity and the low gas velocity rule out current cryovolcanic activity on larger scales, such as that seen on Enceladus<sup>19</sup>. Our observations are compatible with dusty ice

material gently elevated away from surface deposits by sublimation, with no indication for other acceleration mechanisms. The process that has triggered this localized activity—for example, exposure of fresh ice by an impact or removal of a surficial cover by an internally driven explosive event<sup>20–22</sup>—remains unresolved.

Geophysical models suggest that Ceres contains about 25 wt% water in the form of ice and/or bound to minerals, evidenced by its low bulk density<sup>1</sup> of  $2.2 \text{ g cm}^{-3}$ . FC data are consistent with the presence of water ice on Ceres. Unlike the situation at active icy satellites<sup>23,24</sup>, no high-reaching plumes are observed. However, it does seem that there is a significant amount of water ice<sup>25</sup>, as also suggested by evolution models, potentially mixed with other phases, such as sulfates below a surface layer of insulating dark material. The global distribution and the geomorphology of the bright spots, showing lower reflectances than spots in Occator and feature A, but higher reflectances than the background surface, lead us to conclude that these are regions where the water sublimation process has stopped, probably owing to depletion of near-surface water ice resources or insulation by lag deposits. If most bright spots on Ceres represent these formerly active regions, then their global distribution indicates that the amount of near-surface water ice was (and maybe still is) significant, which is in agreement with the observation of relaxed crater morphologies on Ceres<sup>25</sup>. The global distribution of bright spots also suggests the presence of a global subsurface briny water ice layer. The simplest scenario is that the sublimation process of water ice starts after a mixture of ice and salt minerals is exposed by an impact, which penetrates the insulating dark upper crust. The horizontal extent of the near-surface, thin haze cloud corresponds to a zone in which the spectral shape changes from that shown in Fig. 3b (central pit) to those shown in Fig. 3a (outer pit). It reflects the distance-dependent trade-off between supply of ice and dust from the central pit and the final sublimation of the haze and deposition of residual dust.

Ceres is the first identified large body in the main asteroid belt showing not only primitive Solar System material, but also comet-like activity. Our results are consistent with widely distributed subsurface water or ice and water activity in the main asteroid belt, supporting the more recent view of a Solar System with a continuum in composition and ice content between asteroids and comets<sup>26</sup>.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 3 August; accepted 21 September 2015.**

1. Russell, C. T. & Raymond, C. A. (eds) *The Dawn Mission to Minor Planets 4 Vesta and 1 Ceres* (Springer, 2012).
2. McCord, T. B. & Sotin, C. Ceres: evolution and current state. *J. Geophys. Res.* **110**, E05009 (2005).
3. Castillo-Rogez, J. C. & McCord, T. B. Ceres' evolution and present state constrained by shape data. *Icarus* **205**, 443–459 (2010).
4. Küppers, M. *et al.* Localized sources of water vapour on the dwarf planet (1)Ceres. *Nature* **505**, 525–527 (2014).
5. Lebofsky, L. A., Feierberg, M. A., Tokunaga, A. T., Larson, H. P. & Johnson, J. R. The 1.7- to 4.2-micron spectrum of asteroid 1 Ceres: evidence for structural water in clay minerals. *Icarus* **48**, 453–459 (1981).
6. A'Hearn, M. F. & Feldman, P. D. Water vaporization on Ceres. *Icarus* **98**, 54–60 (1992).
7. Rivkin, A. S. *et al.* Hydrogen concentrations on C-class asteroids derived from remote sensing. *Meteorit. Planet. Sci.* **38**, 1383–1398 (2003).
8. Sierks, H. *et al.* in *The Dawn Mission to Minor Planets 4 Vesta and 1 Ceres* (eds Russell, C. T. & Raymond, C. A.) 263–327 (Springer, 2012).
9. Yamamoto, T. & Ashihara, O. Condensation of ice particles in the vicinity of a cometary nucleus. *Astron. Astrophys.* **152**, L17–L20 (1985).
10. Schmidt, J., Brilliantov, N., Spahn, F. & Kempf, S. Slow dust in Enceladus' plume from condensation and wall collision in tiger stripe fractures. *Nature* **451**, 685–688 (2008).
11. Martin, R. G. & Livio, M. On the evolution of the snow line in protoplanetary discs. *Mon. Not. R. Astron. Soc.* **425**, L6–L9 (2012).
12. Reddy, V. *et al.* Color and albedo heterogeneity of Vesta from Dawn. *Science* **336**, 700–704 (2012).
13. Nathues, A. *et al.* Detection of serpentine in exogenic carbonaceous chondrite material on Vesta from Dawn FC data. *Icarus* **239**, 222–237 (2014).

14. Vilas, F. & McFadden, L. A. CCD reflectance spectra of selected asteroids. *Icarus* **100**, 85–94 (1992).
15. De Sanctis, M. C. *et al.* Ammoniated phyllosilicates with a likely outer Solar System origin on (1) Ceres. *Nature* <http://dx.doi.org/10.1038/nature16172> (this issue).
16. Milliken, R. E. & Mustard, J. F. Estimating the water content of hydrated minerals using reflectance spectroscopy: I. Effects of darkening agents and low-albedo materials. *Icarus* **189**, 550–573 (2007).
17. Ehlmann, B. L. *et al.* Geochemical consequences of widespread clay mineral formation in Mars' ancient crust. *Space Sci. Rev.* **174**, 329–364 (2013).
18. Blinn, J. F. Light reflection functions for simulation of clouds and dusty surfaces. *Comput. Graph.* **16**, 21–29 (1982).
19. Hansen, C. J. *et al.* Enceladus' water vapor plume. *Science* **311**, 1422–1425 (2006).
20. Cintala, M. J., Head, J. W. & Parmentier, M. E. Impact heating of H<sub>2</sub>O ice targets: applications to outer planet satellites. *Lunar Planet. Sci.* **XI**, 140–142 (1980).
21. Bowling, T. J., Minton, D. A., Castillo-Rogez, J. C., Johnson, B. C. & Steckloff, J. K. Eroding the hydrosphere of 1 Ceres: water mass loss due to impact induced sublimation. In *Proc. Astrobiology Science Conf. 2015* abstr. 7478, <http://www.hou.usra.edu/meetings/abscicon2015/pdf/7478.pdf> (2015).
22. Williams, N. R., Bell, J. F. III, Christensen, P. R. & Farmer, J. D. Evidence for an explosive origin of central pit craters on Mars. *Icarus* **252**, 175–185 (2015).
23. Soderblom, L. A. *et al.* Triton's geyser-like plumes: discovery and basic characterization. *Science* **250**, 410–415 (1990).
24. Roth, L. *et al.* Transient water vapor at Europa's south pole. *Science* **343**, 171–174 (2014).
25. Schenk, P. *et al.* Impact craters on Ceres: evidence for water-ice mantle? In *European Planetary Science Congress 2015* Vol. 10, <http://meetingorganizer.copernicus.org/EPSC2015/EPSC2015-400.pdf> (2015).
26. Gounelle, M. The asteroid-comet continuum: in search of lost primitivity. *Geosci. World* **7**, 29–34 (2014).

**Acknowledgements** We thank the Dawn operations team for the development, cruise, orbital insertion and operations of the Dawn spacecraft at Ceres. We also thank the FC operations team, especially P. G. Gutierrez-Marques, I. Hall and I. Büttner. The FC project is financially supported by the Max Planck Society and the German Space Agency, DLR.

**Author Contributions** The respective observations were planned by the Dawn science-operations team involving A.N., M.H., M.S., C.A.R., C.T.R. and J.R. A.N., M.H., L.L.C., V.R., T.P., E.A.C., M.R.M.I., D.M.A., N.S. and T.K. contributed to the data analysis. The manuscript was written by A.N., M.H., M.S., L.L.C., V.R., T.P., E.A.C., N.S. and T.K. with reviews and updates by all authors.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.N. (Nathues@mps.mpg.de).



## METHODS

**Crater-based dating of geological units.** To determine an approximate formation age of Occator, we measured the crater size–frequency distribution observed on the continuous ejecta blanket to  $\sim 1$  crater radius from the crater rim following the methodology described in ref. 27. The proximal portion of the ejecta blanket was excluded from the counting area, owing to its hummocky surface texture, to ensure reliable crater identification. The measurements were carried out in a GIS environment (ESRI's ArcGIS) using the CraterTools add-in<sup>28</sup>, which helps to reduce map projection related distortions of crater diameters and counting areas. We used the FC clear filter images 40498, 40499, 40736, 40737, 40753, 40768, 40991 and 41006. The statistical analysis was performed using the CraterStats software<sup>29</sup> including verification for a spatially random distribution of the measured craters<sup>30</sup>.

The cratering model age of  $78 \pm 5$  Ma (that is, the formation age of Occator crater) derived for the ejecta blanket is based on the crater production and chronology functions of ref. 31. They assumed the same projectile distribution as observed on the Moon and a lunar-like time dependence of the projectile flux. These assumptions were also used for asteroid (4) Vesta, where measured crater retention ages are in good agreement with radiometric howardite–eucrite–diogenite (HED) meteorite ages<sup>32</sup>.

We note that measured craters larger than 1.7 km in diameter are superposed by ejecta and, therefore, pre-date the Occator impact. In order to only utilize the smaller crater population superimposed on the ejecta, a resurfacing correction<sup>30</sup> was applied to derive the formation age. The crater size–frequency distribution is shown in a cumulative and differential plot (Extended Data Fig. 3).

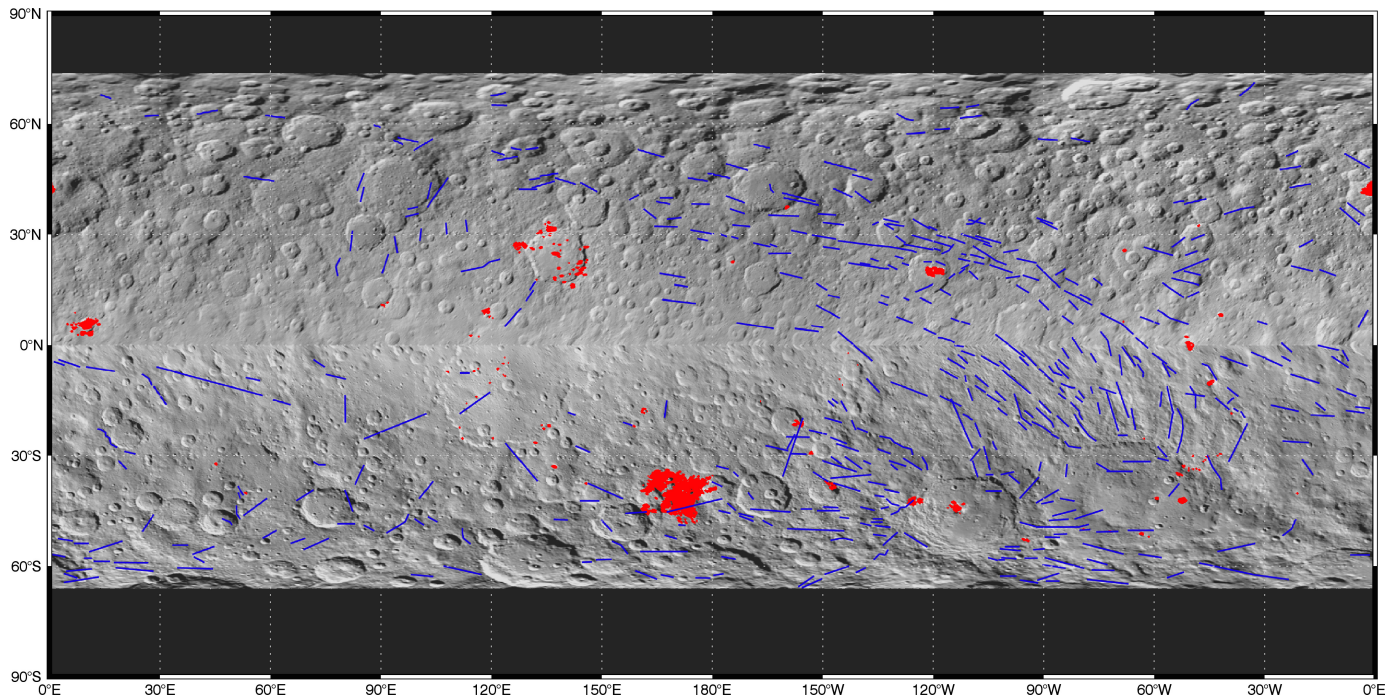
**Code availability.** The software tools CraterTools<sup>28</sup> and CraterStats<sup>29</sup> used to measure and analyse crater size–frequency distributions can be downloaded free of charge at <http://www.geo.fu-berlin.de/en/geol/fachrichtungen/planet/software/index.html>. The Dawn FC images will be made available through the PDS website (<https://pds.jpl.nasa.gov/>).

**Spectral analysis.** We examined the as-measured bright spot spectra and the same spectra after subtraction of the average Ceres spectrum. After subtraction, the bright spot spectra retained their respective red and blue slopes, as well as differences in slope and 'kinks' in the spectra. We compared the salient spectral characteristics to a suite of plausible geological analogues (magnesium sulfates, magnesium carbonates, halide, water ice, hydroxides and clay minerals). Reflectance spectra from RELAB (<http://www.planetary.brown.edu/relabdata/>) and HOSERLab (<http://psf.uwinnipeg.ca/FACILITIES/>; [http://psf.uwinnipeg.ca/Sample\\_Database/](http://psf.uwinnipeg.ca/Sample_Database/)) and ref. 33 were convolved to FC band passes by averaging reflectance values in the laboratory spectra over the FC band passes<sup>34</sup>.

The best matches to the Ceres bright spot spectra in terms of similar spectral shapes and presence of slope changes and kinks longward of  $0.555 \mu\text{m}$  were used to generate areal mixtures involving the average Ceres spectrum. We assumed that this spectrum would be representative of the surrounding 'uncontaminated' materials. In the absence of optical constants for most of these materials, we chose to construct simple mathematical areal mixtures of the average Ceres spectrum with the best end member, matching the spectra at the  $0.653 \mu\text{m}$  FC band pass and weighting each spectrum by an abundance factor. Although mathematical areal mixtures are not ideally representative of how the bright materials are present on Ceres, they do provide a lower limit on the abundance of the bright materials, as bright materials are darkened by opaque phases (such as average Ceres) more in intimate versus areal mixtures. In our spectral comparisons, we focused on using fine-grained ( $< 45 \mu\text{m}$ ) reflectance spectra of bright materials and used only binary mixtures.

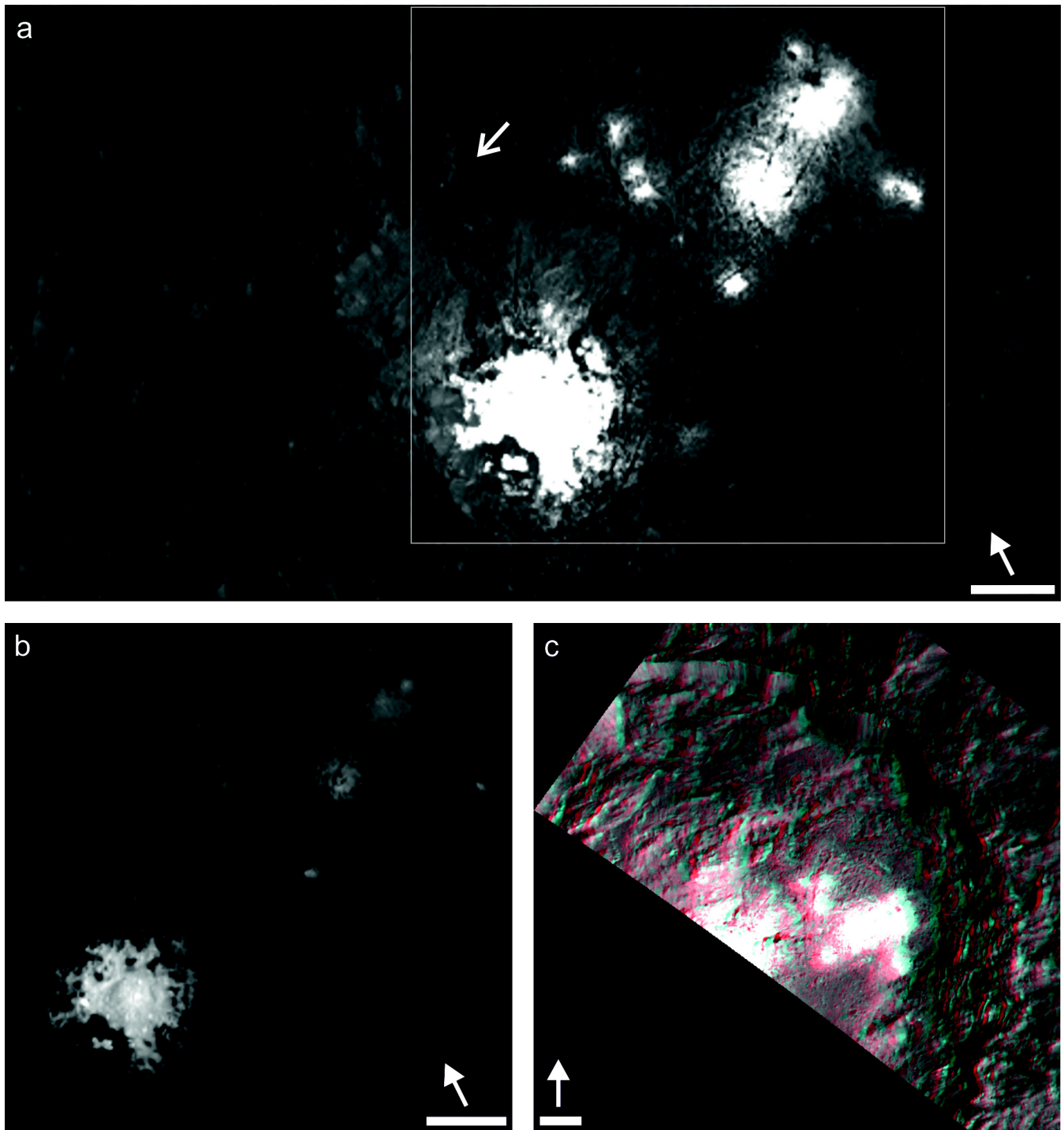
**Sample size.** No statistical methods were used to predetermine sample size.

27. Platz, T., Michael, G. G., Tanaka, K. L., Skinner, J. A. Jr & Fortezzo, C. M. Crater-based dating of geological units on Mars: methods and application for the new global geological map. *Icarus* **225**, 806–827 (2013).
28. Kneissl, T., van Gasselt, S. & Neukum, G. Map-projection-independent crater size-frequency determination in GIS environments — new software tool for ArcGIS. *Planet. Space Sci.* **59**, 1243–1254 (2011).
29. Michael, G. G. & Neukum, G. Planetary surface dating from crater size–frequency distribution measurements: partial resurfacing events and statistical age uncertainty. *Earth Planet. Sci. Lett.* **294**, 223–229 (2010).
30. Michael, G. G., Platz, T., Kneissl, T. & Schmedemann, N. Planetary surface dating from crater size–frequency distribution measurements: spatial randomness and clustering. *Icarus* **218**, 169–177 (2012).
31. Schmedemann, N. *et al.* A preliminary chronology for Ceres. In *Lunar Planet. Sci. Conf.* 46 abstr. 1418, <http://www.hou.usra.edu/meetings/lpsc2015/pdf/1418.pdf> (2015).
32. Schmedemann, N. *et al.* The cratering record, chronology and surface ages of (4) Vesta in comparison to smaller asteroids and the ages of HED meteorites. *Planet. Space Sci.* **103**, 104–130 (2014).
33. Bishop, J. L. *et al.* Spectral properties of Na, Ca-, Mg- and Fe-chlorides and analyses of hydrohalite-bearing samples from Axel Heiberg Island. In *Lunar Planet. Sci. Conf.* 45 abstr. 2145, <http://www.hou.usra.edu/meetings/lpsc2014/pdf/2145.pdf> (2014).
34. Sierks, H. *et al.* The Dawn Framing Camera. *Space Sci. Rev.* **163**, 263–327 (2011).
35. Reddy, V. *et al.* Photometric properties of Ceres from telescopic observations using Dawn Framing Camera color filters. In *Lunar Planet. Sci. Conf.* 46 abstr. 1663, <http://www.hou.usra.edu/meetings/lpsc2015/pdf/1663.pdf> (2015).



**Extended Data Figure 1 | Clear filter mosaic and locations of bright material (red) and global lineaments (blue).** Background map of Ceres is shown in equirectangular projection, combining FC clear filter mosaics from two different sequences of observations during the Rotational Characterization 3 phase at a distance of  $\sim 14,000$  km (RC3). Several mosaics (all sequences from RC3 phase) with different stretches have been

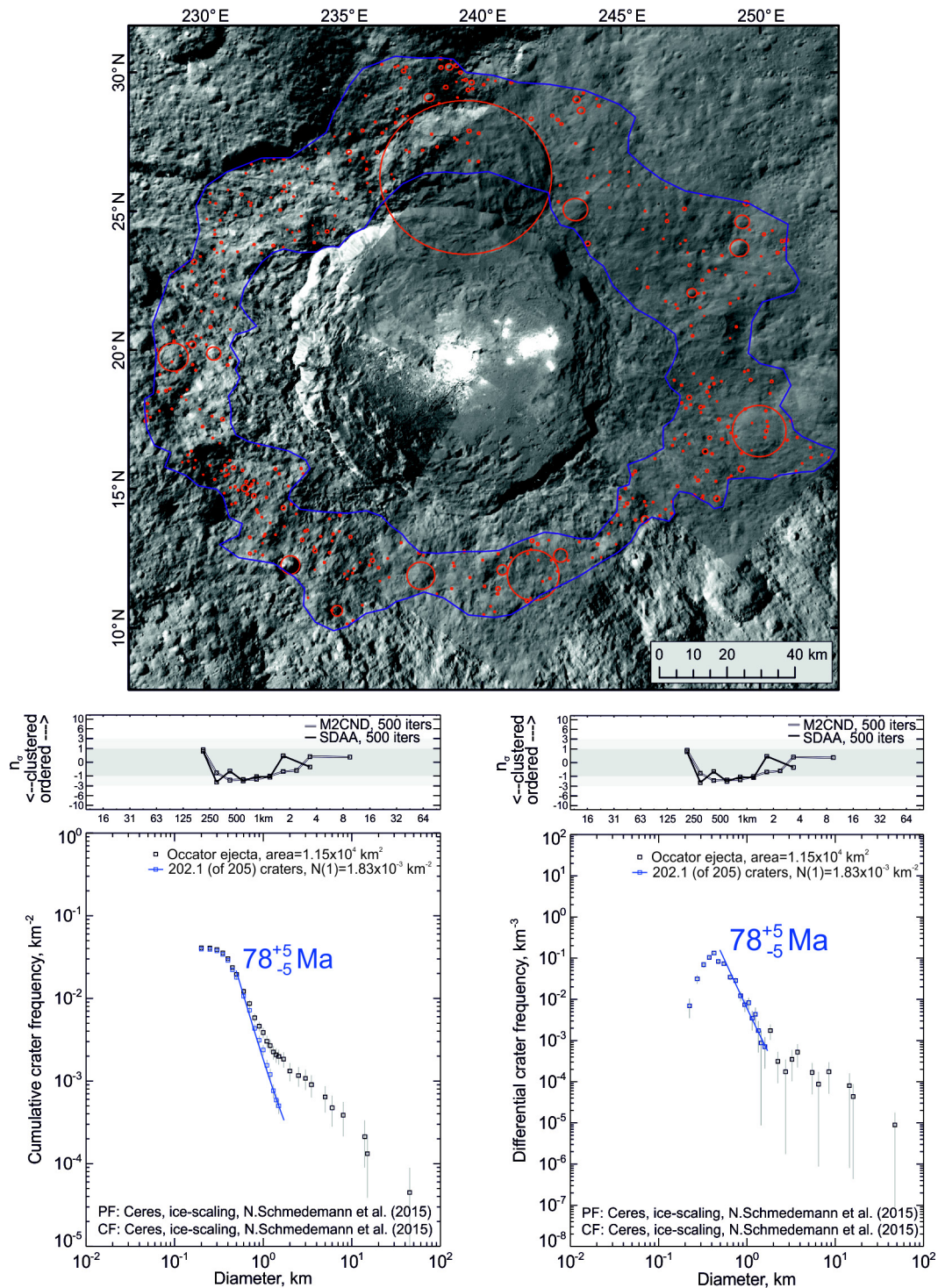
used for the mapping of the linear features. Both fractures and alignment of circular features have been mapped in blue. The former features could correspond to faults, and the latter could be pit crater chains or secondary crater chains. Bright material surfaces are those showing an absolute reflectance larger than 0.037 at  $0.55 \mu\text{m}$ .



**Extended Data Figure 2 | Interior of Occator crater.** **a**, Image shows a scene of the centre of Occator crater from High Altitude Mapping Orbit (FC image 40752, resolution  $\sim 140$  m per pixel) revealing curvilinear depressions and a smooth pond-like feature (arrow at upper centre) at one of its ends. A possible explanation of the last feature is former short-term liquid-flow material. **b**, Central and peripheral bright spots within Occator crater. A nonlinear stretch is applied to enhance the interior structure of the bright spot. **c**, 3D anaglyph (red–cyan glasses are required to see this

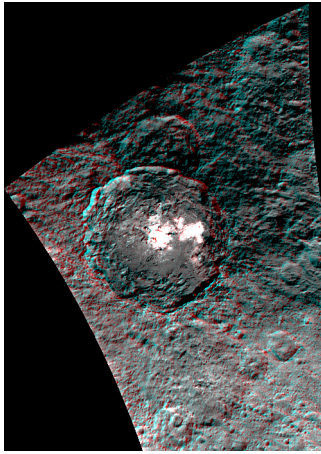
scene in 3D) showing the northeastern portion of Occator crater. A flow lobe with well-defined margins at the image's centre is clearly visible. Note that the extent of secondary bright spot occurrences is confined to the extent of the flow lobe. Extensive slumping of wall material is observed in the northern portion of the crater. The anaglyph (images 40736 and 40752) has a vertical exaggeration of 1.9 if viewed at a distance of 50 cm. Arrows above scale bars point towards north; scale bars are 7 km long.



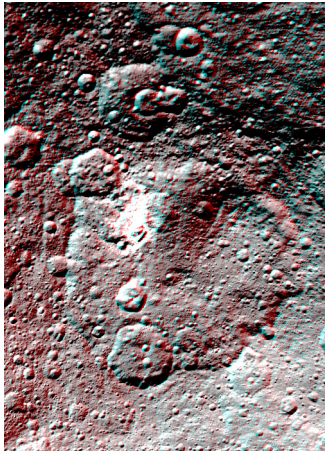


**Extended Data Figure 3 | Crater counting area, superposed impact craters and model age.** The upper panel shows the crater counting area (blue outline) and measured impact craters (red outlines) on the ejecta blanket of Occator crater (FC images 40498, 40499, 40736, 40737, 40753, 40768, 40991 and 41006). On the basis of the mapped crater size–frequency distribution, we derived a model age of  $78 \pm 5$  Ma for the formation of Occator crater (lower panels) shown in cumulative (left) and differential (right) plots. To model an absolute age we used the production function (PF) and chronology function (CF) of ref. 31 and fitted the isochron to the crater diameter range 0.5–1.7 km. Error bars are  $\pm$  the square-root of the cumulative number of craters per bin

divided by the counting area—the largest bins contain fewer craters and therefore have the largest error bars. The reference value  $N(1)$  denotes the cumulative number of craters larger than 1 km in diameter per square kilometre. Upper insets show results from a Monte-Carlo-based randomness analysis for the measured crater population using the mean second-closest neighbour distance ('M2CND') and the standard deviation of adjacent area ('SDAA') methods<sup>30</sup>. Standard deviations  $n_\sigma$  are plotted above and below the Monte-Carlo-derived mean. Details about the methods are given in ref. 30. Here, the observed crater population plots within the  $3\sigma$  range indicating a random distribution.

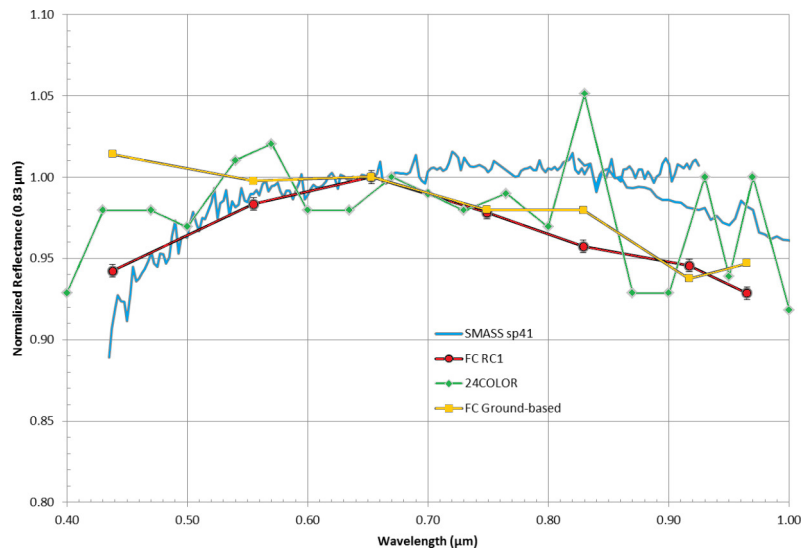


**Extended Data Figure 4 | 3D anaglyph of Occator crater.** The interior of the 90.5-km-diameter crater is characterized by an abundance of terraces and a smooth inner floor surface. The inner crater rim appears scalloped in places. The central pit with the brightest spot on Ceres is partially rimmed. Occator ejecta extend up to one crater diameter outwards, partially or completely burying pre-existing terrain and impact craters. The anaglyph is composed of FC images 37674 and 37666. Vertical exaggeration is approximately 4.5 if viewed at a distance of 50 cm.



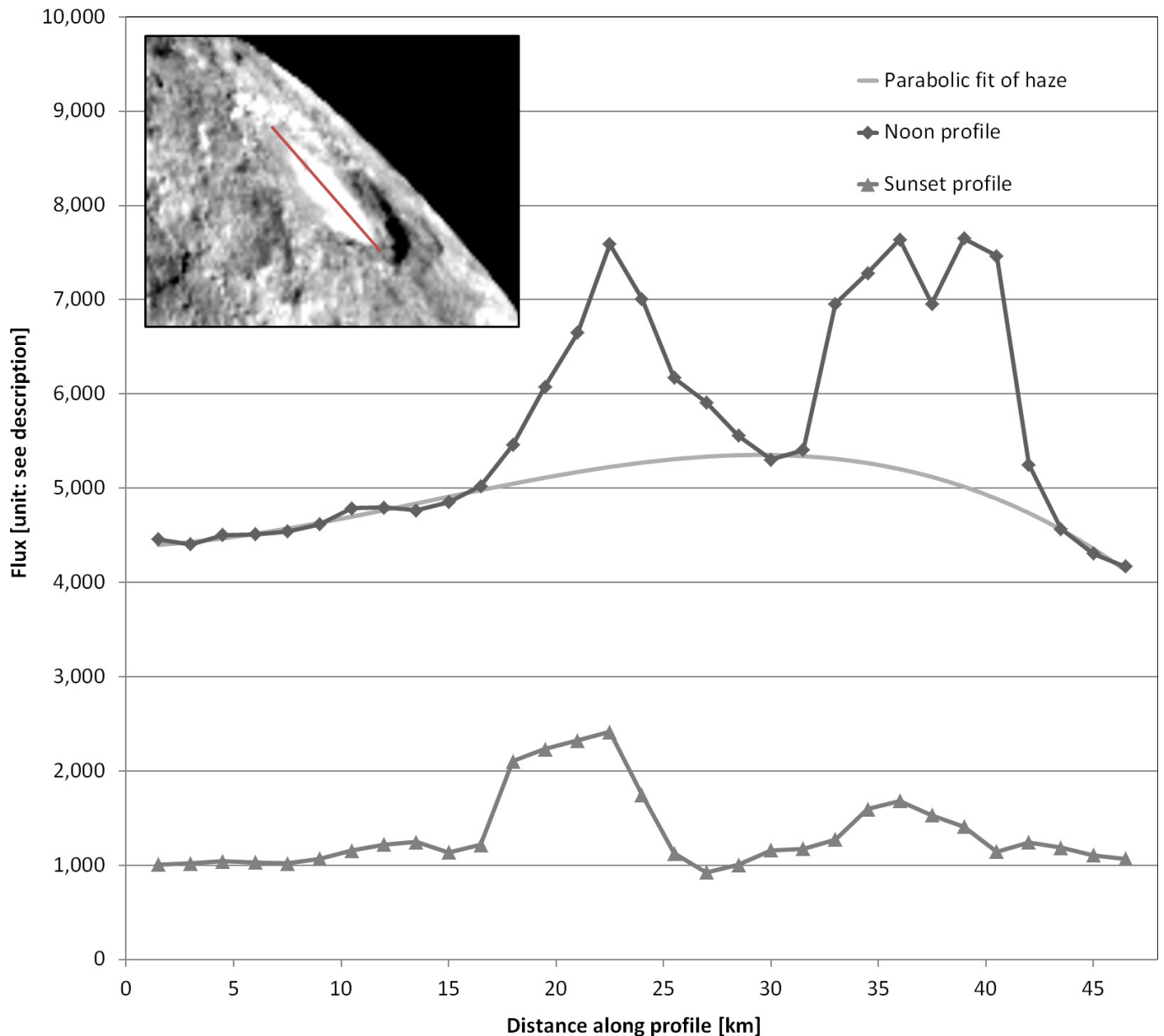
**Extended Data Figure 5 | 3D anaglyph of an unnamed crater hosting feature A.** The large crater accommodating feature A is degraded and marked by a low-relief crater rim partially eroded by subsequent impact craters. Its interior exhibits a densely cratered floor and hosts an asymmetric dome. The bright material is exposed within a 10-km-diameter crater on the wall of the larger, degraded crater and its vicinity. 3D anaglyph is composed of FC images 38409 and 38407. Vertical exaggeration is approximately 3.0 if viewed at a distance of 50 cm.





**Extended Data Figure 6 | Average FC colour spectrum ('FC RC1') of Ceres and ground-based spectra.** Ground-based spectra ('SMASS sp41', '24COLOR' and 'FC Ground-based') of Ceres are presented for comparison. Ceres spectra, obtained by ground-based telescopes, in general show a large variety. The ground-based average colour spectrum of Ceres using FC spare filters<sup>35</sup> is, except for filter 0.44  $\mu\text{m}$ , in good

agreement with the in-flight spectrum. We note that the ground-based spectra are obtained under different viewing geometries and are not photometrically corrected compared to the in-flight data. Ground-based spectra 'SMASS sp41' and '24COLOR' are available at <http://sbn.psi.edu/pds/archive/asteroids.html>. Spectrum 'FC Ground-based' is presented in ref. 35.



**Extended Data Figure 7 | Haze cloud intensity profile.** Two flux profiles from FC clear filter images across the central part of Occator crater (indicated by the red line; see inset). Units of flux are counts per pixel ( $\sim 1.5 \text{ km} \times 1.5 \text{ km}$ ). Each profile crosses the centres of the brightest and second brightest spots (see inset, noon image 37113). Two images were selected, the first (37113) showing the scene at low incidence and high emission angles, that is, an oblique view near local noon ('Noon profile'). The second image (36681) was obtained at high angles in both emission

and incidence, showing an oblique view near local sunset ('Sunset profile'). While the 'Noon profile' shows a signal enhancement centred between the peaks, the 'Sunset profile' does not show this. The enhancement is best described by a parabolic fit of the background, ignoring the signal peaks ('Parabolic fit of haze'). The shape of this fit is consistent with a diffuse light scattering component in Occator crater at maximum insolation, observable at very oblique views, which increase the length of the light path through the layer, resembling the phenomena associated with a haze layer.

# Ammoniated phyllosilicates with a likely outer Solar System origin on (1) Ceres

M. C. De Sanctis<sup>1</sup>, E. Ammannito<sup>1,2</sup>, A. Raponi<sup>1</sup>, S. Marchi<sup>1,3</sup>, T. B. McCord<sup>4</sup>, H. Y. McSween<sup>5</sup>, F. Capaccioni<sup>1</sup>, M. T. Capria<sup>1</sup>, F. G. Carrozzo<sup>1</sup>, M. Ciarniello<sup>1</sup>, A. Longobardo<sup>1</sup>, F. Tosi<sup>1</sup>, S. Fonte<sup>1</sup>, M. Formisano<sup>1</sup>, A. Frigeri<sup>1</sup>, M. Giardino<sup>1</sup>, G. Magni<sup>1</sup>, E. Palomba<sup>1</sup>, D. Turrini<sup>1</sup>, F. Zamboni<sup>1</sup>, J.-P. Combe<sup>4</sup>, W. Feldman<sup>6</sup>, R. Jaumann<sup>7</sup>, L. A. McFadden<sup>8</sup>, C. M. Pieters<sup>9</sup>, T. Prettyman<sup>6</sup>, M. Toplis<sup>10</sup>, C. A. Raymond<sup>11</sup> & C. T. Russell<sup>2</sup>

Studies of the dwarf planet (1) Ceres using ground-based and orbiting telescopes have concluded that its closest meteoritic analogues are the volatile-rich CI and CM carbonaceous chondrites<sup>1,2</sup>. Water in clay minerals<sup>3</sup>, ammoniated phyllosilicates<sup>4</sup>, or a mixture of Mg(OH)<sub>2</sub> (brucite), Mg<sub>2</sub>CO<sub>3</sub> and iron-rich serpentine<sup>5,6</sup> have all been proposed to exist on the surface. In particular, brucite has been suggested from analysis of the mid-infrared spectrum of Ceres<sup>6</sup>. But the lack of spectral data across telluric absorption bands in the wavelength region 2.5 to 2.9 micrometres—where the OH stretching vibration and the H<sub>2</sub>O bending overtone are found—has precluded definitive identifications. In addition, water vapour around Ceres has recently been reported<sup>7</sup>, possibly originating from localized sources. Here we report spectra of Ceres from 0.4 to 5 micrometres acquired at distances from ~82,000 to 4,300 kilometres from the surface. Our measurements indicate widespread ammoniated phyllosilicates across the surface, but no detectable water ice. Ammonia, accreted either as organic matter or as ice, may have reacted with phyllosilicates on Ceres during differentiation. This suggests that material from the outer Solar System was incorporated into Ceres, either during its formation at great heliocentric distance or by incorporation of material transported into the main asteroid belt.

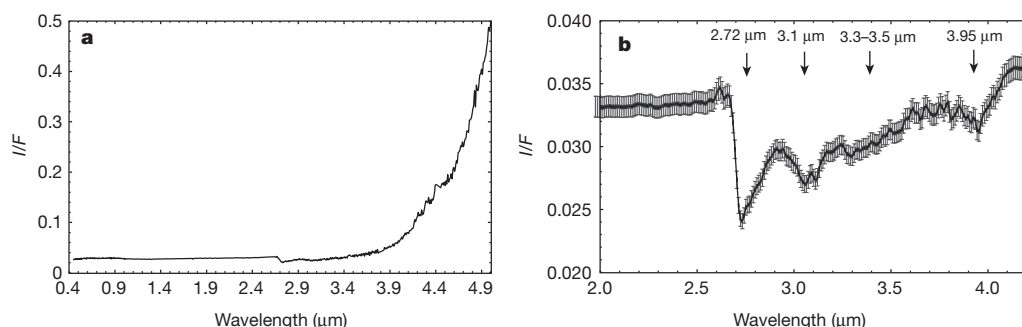
The Visible-Infrared Mapping Spectrometer VIR<sup>8</sup> on the Dawn spacecraft<sup>9</sup> obtained multi-spectral images of Ceres' surface in the spectral range 0.4–5 µm with an instantaneous field of view of 250 µrad (Fig. 1). This range includes the 2.6–2.9 µm spectral region precluded

from telescopic measurements owing to the atmospheric absorptions. As such, Dawn data provide new constraints on Ceres' surface composition.

The average spectrum of Ceres' surface measured by VIR shows a low level of reflectance combined with prominent thermal emission at longer wavelengths, indicative of a very dark, warm surface. At the spatial scale of ~11 km per pixel, daytime surface temperatures of Ceres span the range 180–240 K at a heliocentric distance of 2.86 AU, with maximum temperatures inferred in the equatorial region, corresponding to the minimum solar incidence angle. This range of temperatures is generally incompatible with a widespread occurrence of water ice on the surface, even if very small localized occurrences of water ice cannot be excluded<sup>10</sup>.

The geometric albedo obtained from VIR is  $0.088 \pm 0.006$  at 0.55 µm, derived from Hapke modelling of the average surface photometric properties of the body<sup>11</sup>. The measured value of the albedo agrees with Hubble Space Telescope observations in the F555W filter<sup>12</sup>. Given the low albedo of the surface, single scattering dominates the reflectance. The thermally corrected reflectance spectrum of Ceres (Fig. 1b) shows that the 2.6–4.2 µm wavelength region is characterized by a broad asymmetric feature, characteristic of H<sub>2</sub>O/OH bearing materials. Within this broad absorption are several distinct absorption bands at 2.72 µm, 3.05–3.1 µm, 3.3–3.5 µm and 3.95 µm.

Laboratory spectra of CM and CI carbonaceous chondrites measured under anhydrous conditions<sup>13</sup> show the same prominent 2.7 µm OH absorption band as the Ceres spectrum, suggesting



**Figure 1 | Average spectrum of Ceres.** **a**, Average spectrum of Ceres from 0.4 µm to 5 µm. No smoothing has been applied to the spectrum and some instrument artefacts are still present, especially in the thermal range (4.2–4.6 µm). Error bars were calculated taking into account a mean absolute deviation of the calibration uncertainties along the 256 samples,

but are not shown for clarity; they are of the same order as those in **b**. **b**, Expanded thermal-removed average Ceres spectrum in the 2–4.2 µm range. Main bands are indicated by arrows.  $I/F$  is the spectral radiance divided by the solar spectral irradiance.

<sup>1</sup>Istituto di Astrofisica e Planetologia Spaziali, INAF, Via del Fosso del Cavaliere 100, 00133 Roma, Italy. <sup>2</sup>Institute of Geophysics and Planetary Physics, University of California, Los Angeles,

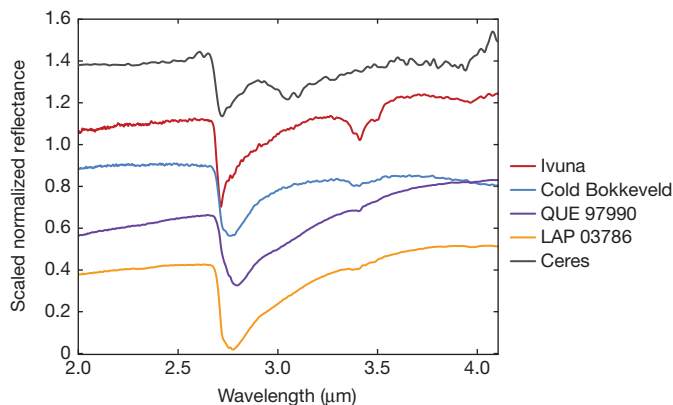
California 90095-1567, USA. <sup>3</sup>Southwest Research Institute, 1050 Walnut Street, Boulder, Colorado 80302, USA. <sup>4</sup>Bear Fight Institute, 22 Fiddler's Road, PO Box 667, Winthrop, Washington

98862, USA. <sup>5</sup>Department of Earth and Planetary Sciences, University of Tennessee, Knoxville, Tennessee 37996-1410, USA. <sup>6</sup>Planetary Science Institute, Tucson, Arizona 85719-2395, USA.

<sup>7</sup>Institute of Planetary Research, German Aerospace Center (DLR), Rutherfordstrasse 2, 12489 Berlin, Germany. <sup>8</sup>NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA.

<sup>9</sup>Department of Earth, Environmental, and Planetary Sciences, Brown University, Providence, Rhode Island 02912, USA. <sup>10</sup>Institut de Recherche d'Astrophysique et Planétologie, Observatoire Midi Pyrénées, Université Paul Sabatier, 14 Avenue E. Belin, 31400 Toulouse, France. <sup>11</sup>Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA.



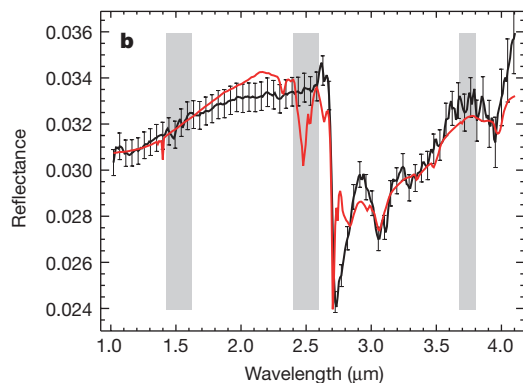
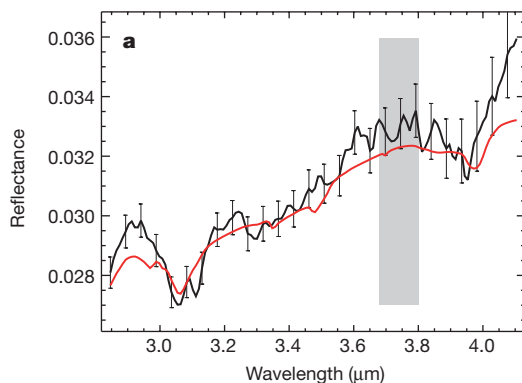


**Figure 2 | Spectrum of Ceres compared with spectra of carbonaceous chondrites.** Shown are normalized reflectance spectra of CM/CI carbonaceous chondrites (Ivuna, Cold Bokkeveld, QUE 97990 and LAP 03786) measured under dry conditions<sup>13</sup> compared with Ceres' spectrum acquired by VIR. The spectra have been offset for clarity. We thank D. Takir for the spectra of carbonaceous chondrites.

some mineralogical similarity, consistent with the idea that these materials are relevant meteoritic analogues for Ceres<sup>1,2</sup>. However, these chondrites are not a good match with Ceres at other wavelengths (Fig. 2). Even if the spectrum of Ceres is broadly similar to that of CM chondrites, distinct spectral features indicate variations in mineral proportions that may reflect differences in alteration pathways and/or formation conditions. For instance, the  $\sim 4.0\text{ }\mu\text{m}$  absorption is stronger in the Ceres spectrum than in most CM/CI spectra, whereas the  $\sim 3.4\text{ }\mu\text{m}$  feature is usually stronger in CM/CIs. Moreover, CM and CI chondrites lack Ceres' distinctive signature at  $3.05\text{--}3.1\text{ }\mu\text{m}$ .

The  $3.05\text{--}3.1\text{ }\mu\text{m}$  band dominates Ceres' ground-based spectra, and was also observed on two other asteroids<sup>14</sup>. This signature has been attributed to a variety of different phases including water ice, hydrated or  $\text{NH}_4$ -bearing clays and brucite<sup>3–6</sup> (Extended Data Fig. 1). Brucite, in particular, clearly shows this narrow characteristic absorption<sup>15</sup>. Ammoniated mineral species, including  $\text{NH}_4$ -bearing annite and  $\text{NH}_4$ -montmorillonite, have an absorption feature near  $3.06\text{ }\mu\text{m}$ , which could account for this band.

VIR observations show a strong and narrow absorption centred at  $2.72\text{--}2.73\text{ }\mu\text{m}$ . This characteristic feature is distinctive for OH-bearing minerals (Extended Data Fig. 1).  $\text{H}_2\text{O}$ -bearing phases, however, show a much broader absorption band that is a poor match for the Ceres spectrum. Water ice does not fit the observed spectrum, consistent with its instability at the maximum surface temperature. OH stretching vibrations occur in the  $2.7\text{--}2.85\text{ }\mu\text{m}$  range for phyllosilicates<sup>16</sup>, with band centres at different wavelengths for different species.



**Figure 3 | Spectral fits of the spectrum of Ceres.** **a**, Results of the spectral fitting model (red curve) using brucite, Mg-carbonate, dark component and cronstedtite<sup>6</sup>; **b**, the same model (red curve) is shown over an extended spectral range ( $1\text{--}4.2\text{ }\mu\text{m}$ ). Vertical grey bars represent regions

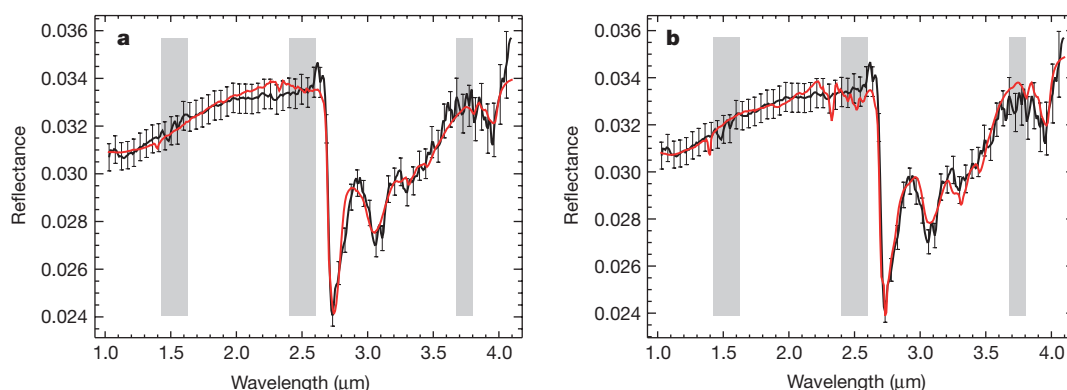
Ceres is rich in absorption bands in the  $3\text{--}4\text{ }\mu\text{m}$  region, but the spectral range  $1.0\text{--}2.5\text{ }\mu\text{m}$  lacks distinctive features. Conversely, some terrestrial  $\text{H}_2\text{O}$ - and OH-bearing phases show numerous diagnostic bands over this spectral range. For example, brucite has a prominent  $2.46\text{ }\mu\text{m}$  band that is absent in the Ceres spectra (Extended Data Fig. 1). Also, the vibrations due to the OH overtone at about  $1.4\text{ }\mu\text{m}$  and the bands due to H–O–H combination stretching and bending vibrations in the  $1.9\text{ }\mu\text{m}$  region are absent in the Ceres spectrum. The combination of phyllosilicates with other species, however, can mask the presence of such bands.

In order to model Ceres' surface spectrum, we solved the radiative transfer equation in a particulate medium, using the Hapke model<sup>17</sup>. We investigated several different combinations of mineralogical phases (Methods and Extended Data Tables 1 and 2). We started with a simple combination of water ice with a dark component (amorphous carbon or magnetite), but the presence of ice gives rise to clear absorption bands at  $1.5\text{ }\mu\text{m}$  and  $2\text{ }\mu\text{m}$  that are not present in our data (Extended Data Fig. 2). Also, the  $3\text{ }\mu\text{m}$  ice band centre does not fit the Ceres data. Magnetite (or carbon) in the fits represents a highly absorbing species whose specific character is not defined in this analysis. However, the mixture with magnetite fits the spectral slope between  $1\text{ }\mu\text{m}$  and  $2.6\text{ }\mu\text{m}$  better. For this reason we choose to use magnetite in the other mixtures.

While the mixture of brucite,  $\text{MgCO}_3$ , a dark component, and Fe-serpentine (cronstedtite)<sup>6</sup> is able to match the Ceres spectrum in the  $2.9\text{--}4.2\text{ }\mu\text{m}$  range, it fails to fit the intense  $2.7\text{ }\mu\text{m}$  band (Fig. 3). Also, mixtures of water ice with CM or CI carbonaceous chondrite do not fit the Ceres spectrum (Extended Data Fig. 3 and Extended Data Table 2). The fit is also poor in the case of adding brucite and water to carbonaceous chondrites (Extended Data Fig. 3 and Extended Data Table 2). The prominent  $2.7\text{ }\mu\text{m}$  band can be fitted by a mixture of phyllosilicates such as Mg-serpentine (antigorite) or tochilinite (structurally interlayered brucite and iron sulphide), with Fe-serpentine (cronstedtite), carbonate (dolomite) and magnetite (Extended Data Fig. 4a, b). The fit with antigorite is slightly better in reproducing the depth and centre of the  $2.7\text{ }\mu\text{m}$  band; thus for the following fits we use antigorite to reproduce the  $2.7\text{ }\mu\text{m}$  feature. However, the fit does not fully reproduce the absorption at  $3.05\text{--}3.1\text{ }\mu\text{m}$ . By analogy with CM chondrites, however, tochilinite is more likely than antigorite to be in equilibrium with cronstedtite and thus may be a more plausible phase<sup>18</sup>. Addition of brucite to the mixture of phyllosilicates, carbonates and magnetite described above slightly improves the fits at  $3.05\text{--}3.1\text{ }\mu\text{m}$ , but the fits are not satisfactory at shorter wavelengths (Extended Data Fig. 4c, d and Extended Data Table 2).

Other species that have clear absorptions at  $3.05\text{--}3.1\text{ }\mu\text{m}$  are ammoniated-bearing minerals, which also show absorptions at  $3.3\text{ }\mu\text{m}$  (ref. 19). The best fit of Ceres' spectrum over the infrared range is consistently obtained by adding ammoniated phyllosilicates to a dark material

of the detector affected by the filter junctions. Ceres' spectrum is in black. Error bars are calculated taking into account a mean absolute deviation of the calibration uncertainties along the 256 samples.



**Figure 4 | Spectral fits of the spectrum of Ceres with ammonia-bearing species.** **a**, Results of the spectral fitting model (red curve) using  $\text{NH}_4$ -montmorillonite, antigorite, Mg-carbonate, and dark component;

**b**, results of the spectral fitting model (red curve) using  $\text{NH}_4$ -annite, antigorite, Mg-carbonate, and dark component. Grey bars as in Fig. 3. Ceres' spectrum is in black. Error bars as in Fig. 3.

(magnetite), antigorite and carbonate (Fig. 4). Mixtures with different ammoniated clays (Fig. 4) can match the Ceres spectrum, but the specific clay mineral cannot be determined. Carbonates are always needed in the fit but the specific carbonate mineral is not fully constrained because dolomite, magnesite and calcite produce equivalent matches to the observed Ceres spectrum (Extended Data Fig. 5 and Extended Data Table 2).

Ammoniated clays on Ceres have been suggested previously<sup>4</sup>, possibly formed by reaction with ammonia ice. Recent observation of substantial amounts of ammonia contained in organic compounds in some CI, CM and CR chondrites<sup>20</sup> suggests another mechanism for the formation of ammoniated clays on Ceres. The consistent  $^{15}\text{N}$ -enrichment of carbonaceous chondrite organic matter may indicate that enrichment occurred in the solar nebula before star formation<sup>20</sup>, although ammonia was also likely to have been processed into more complex molecules in the protoplanetary disk and parent bodies. Modest heating to 300–400 °C of  $\text{NH}_4$ -bearing organic matter in carbonaceous chondrites releases ammonia<sup>20</sup>, and a similar process on Ceres could have plausibly produced ammoniated phyllosilicates. The ability of smectite, especially montmorillonite, to incorporate ammonia by cation exchange is well documented<sup>21</sup>.

Ceres' surface enrichment in ammoniated phyllosilicates and relatively low bulk density pose significant challenges to understanding its origin and formation. One view holds that Ceres and most C-type asteroids could have formed close to their current locations, inferred to be near the water snow line<sup>22</sup>. However, CI/CM meteorites have bulk water contents of 15%–20% at most, whereas Ceres is inferred to be ~30% water by mass, on the basis of its bulk density. This may indicate that Ceres retained more volatiles than objects represented in the meteorite collection, or that it accreted from more volatile-rich material.

We have shown here that the altered CM and CI carbonaceous chondrites do not offer a good spectral match for Ceres, nor do all carbonaceous chondrites show evidence of exposure to significant amounts of ammonia. Therefore, it seems unlikely that a substantial population of ammonia-rich planetesimals could have formed locally at Ceres' current location in the main belt, even if the water snow line episodically migrated about 1 AU during the evolution of the protoplanetary disk<sup>22–24</sup>. However, it is possible that Ceres grew close to its present position by accreting pebble-sized objects, some of which could have drifted inward from larger heliocentric distances<sup>25</sup> where ammonia was stable. If the source of  $\text{NH}_3$  for ammoniated phyllosilicate was ammonia-bearing organic matter, Ceres could have accreted organics originally formed by irradiation of  $\text{N}_2$  ices condensed near or beyond the orbit of Neptune.

Alternatively, Ceres may itself have formed further out in the Solar System, presumably in the trans-Neptunian disk, before being subsequently implanted in the main belt<sup>26</sup>. This view is corroborated by the presence of ammonia ice on other large trans-Neptunian objects, such

as Orcus and Charon<sup>27,28</sup>. The implantation of Ceres in the main belt could have taken place during a migratory phase of the giant planets, either during their growth in the protoplanetary disk<sup>29</sup>, or at a later time as a result of an orbital instability<sup>30</sup>.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 28 July; accepted 29 October 2015.**

- Chapman, C. R. & Salisbury, J. W. Comparisons of meteorite and asteroid spectral reflectivities. *Icarus* **19**, 507–522 (1973).
- McCord, T. & Gaffey, M. J. Asteroids: surface composition from reflectance spectroscopy. *Science* **186**, 352–355 (1974).
- Lebofsky, L., Feierberg, M., Tokunaga, A., Larson, H. & Johnson, J. The 1.7–4.2  $\mu\text{m}$  spectrum of asteroid 1 Ceres: evidence for structural water in clay minerals. *Icarus* **48**, 453–459 (1981).
- King, T., Clark, R., Calvin, W., Sherman, D. & Brown, R. Evidence for ammonium-bearing minerals on Ceres. *Science* **255**, 1551–1553 (1992).
- Rivkin, A. S., Volquardsen, E. L. & Clark, B. E. The surface composition of Ceres: discovery of carbonates and iron-rich clays. *Icarus* **185**, 563–567 (2006).
- Milliken, R. E. & Rivkin, A. S. Brucite and carbonate assemblages from altered olivine-rich materials on Ceres. *Nature Geosci.* **2**, 258–261 (2009).
- Küppers, M. *et al.* Localized sources of water vapour on the dwarf planet (1) Ceres. *Nature* **505**, 525–527 (2014).
- De Sanctis, M. C. *et al.* The VIR spectrometer. *Space Sci. Rev.* **163**, 329–369 (2011).
- Russell, C. T. & Raymond, C. A. The Dawn Mission to Vesta and Ceres. *Space Sci. Rev.* **163**, 3–23 (2011).
- Nathues, A. *et al.* Sublimation in bright spots on (1) Ceres. *Nature* <http://dx.doi.org/10.1038/nature15754> (this issue).
- Helfenstein, P. & Veeverka, J. in *Asteroids II* (eds Binzel, R. P., Gehrels, T. & Matthews, M. S.) 557–593 (Univ. Arizona Press, 1989).
- Li, J.-Y. *et al.* Photometric analysis of 1 Ceres and surface mapping from HST observations. *Icarus* **182**, 143–160 (2006).
- Takir, D. *et al.* Nature and degree of aqueous alteration in CM and CI carbonaceous chondrites. *Meteorit. Planet. Sci.* **48**, 1618–1637 (2013).
- Takir, D. & Emery, J. P. Outer main belt asteroids: identification and distribution of four 3- $\mu\text{m}$  spectral groups. *Icarus* **219**, 641–654 (2012).
- Beck, P., Schmitt, B., Cloutis, E. A. & Vernazza, P. Low-temperature reflectance spectra of brucite and the primitive surface of 1-Ceres? *Icarus* **257**, 471–476 (2015).
- Bishop, J. L. *et al.* Reflectance and emission spectroscopy study of four groups of phyllosilicates: smectites, kaolinite-serpentines, chlorites and micas. *Clay Miner.* **43**, 35–54 (2008).
- Hapke, B. *Theory of Reflectance and Emission Spectroscopy* (Cambridge Univ. Press, 2012).
- Howard, K. T., Alexander, C. M. O'D., Schrader, D. L. & Dyl, K. A. Classification of hydrous meteorites (CR, CM and C2 ungrouped) by phyllosilicate fraction: PSD-XRD modal mineralogy and planetesimal environments. *Geochim. Cosmochim. Acta* **149**, 206–222 (2015).
- Bishop, J. L., Banin, A., Mancinelli, R. L. & Klovstad, M. L. Detection of soluble and fixed  $\text{NH}_4^+$  in clay minerals by DTA and IR reflectance spectroscopy: a potential tool for planetary surface exploration. *Planet. Space Sci.* **50**, 11–19 (2002).
- Pizzarello, S. & Williams, L. B. Ammonia in the early Solar System: an account from carbonaceous meteorites. *Astrophys. J.* **749**, 161–167 (2012).

21. Borden, D. & Giese, R. F. Baseline studies of the Clay Minerals Society source clays: cation exchange capacity measurements by the ammonia-electrode method. *Clays Clay Miner.* **49**, 444–445 (2001).
22. Martin, R. G. & Livio, M. On the evolution of the snow line in protoplanetary discs. *Mon. Not. R. Astron. Soc. Lett.* **425**, L6–L9 (2012).
23. Lodders, K. Solar System abundances and condensation temperatures of the elements. *Astrophys. J.* **591**, 1220–1247 (2003).
24. Bitsch, B., Johansen, A., Lambrechts, M. & Morbidelli, A. The structure of protoplanetary discs around evolving young stars. *Astron. Astrophys.* **575**, A28 (2015).
25. Johansen, A., Mac Low, M.-M., Lacerda, P. & Bizzarro, M. Growth of asteroids, planetary embryos, and Kuiper belt objects by chondrule accretion. *Sci. Adv.* **1**, 1500109 (2015).
26. McKinnon, W. B. Where did Ceres accrete? In *Proc. Conf. Asteroids, Comets, Meteors 2012* abstr. 6475, <http://www.lpi.usra.edu/meetings/acm2012/pdf/6475.pdf> (2012).
27. DeMeo, F. E. *et al.* Spectral variability of Charon's 2.21- $\mu\text{m}$  feature. *Icarus* **246**, 213–219 (2015).
28. Barucci, M. A. *et al.* Surface composition and temperature of the TNO Orcus. *Astron. Astrophys.* **479**, L13–L16 (2008).
29. Walsh, K. J., Morbidelli, A., Raymond, S. N., O'Brien, D. P. & Mandell, A. M. A low mass for Mars from Jupiter's early gas-driven migration. *Nature* **475**, 206–209 (2011).
30. Levison, H. *et al.* Contamination of the asteroid belt by primordial trans-Neptunian objects. *Nature* **460**, 364–366 (2009).

**Acknowledgements** We thank the following institutions and agencies, which supported this work: the Italian Space Agency (ASI), the National Aeronautic and Space Administration (NASA, USA) and the Deutsches Zentrum für Luft- und Raumfahrt (DLR, Germany). The VIR was funded and coordinated by the Italian Space Agency and built by SELEX ES, with the scientific leadership of the Institute for Space Astrophysics and Planetology, Italian National Institute for Astrophysics, Italy, and is operated by the Institute for Space Astrophysics and Planetology, Rome, Italy. A portion of this work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract to NASA. We thank J. L. Bishop and D. Takir for reviews, and D. Takir for providing spectra of carbonaceous chondrites plotted in Fig. 2.

**Author Contributions** M.C.D.S., A.R., E.A. and G.C. performed data analysis and calibration. M.C. provided optical constants from reflectance spectra. F.T. provided the surface temperatures. M.C.D.S., S.M., H.Y.M., T.B.M. and C.M.P. contributed to the interpretation of the data. All authors contributed to the discussion of the results and to writing the paper.

**Author Information** The VIR calibrated data will be made available through the PDS Small Bodies Node website (<http://sbn.pds.nasa.gov/>). Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.C.D.S. ([mariacristina.desanctis@iaps.inaf.it](mailto:mariacristina.desanctis@iaps.inaf.it)).



## METHODS

**Spectral modelling.** The surface spectrum of Ceres reported in the main text is derived as a median of ~2,500 spectra acquired during the first phase of the mission, under standard viewing geometry (incidence angle 30°, emission angle 0°). The resulting spectrum has been corrected for thermal emission by removing a Planck function produced by the best fit with the measured spectra. We model the Ceres average spectrum as formed by an intimate mixture of different end-members, starting from their optical constants. We take advantage of Hapke theory<sup>17</sup> which characterizes light scattering in particulate media. The optical constants, when not available from literature, are derived from reflectance spectra as in ref. 31. The end-members treated in this work are listed in Extended Data Table 1. The best fit between the Ceres spectrum and the models are obtained by means of a least squares optimization algorithm as in ref. 32, where the free parameters are the abundances and grain sizes of each end-member (abundances, grain sizes and  $\chi^2$  reported in Extended Data Table 2).

**Laboratory spectra of minerals.** The carbonaceous chondrite laboratory reflectance spectra used as end-members presented were measured at the Reflectance Experiment Laboratory (RELAB) at Brown University and obtained through the RELAB online database. Ivuna and Murchison samples were heated in a vacuum cell at different temperatures for one week<sup>33,34</sup>. The sample lost a large part of its 3  $\mu\text{m}$  water band by heating at 300 °C. MAC87300 is a candidate for thermally metamorphosed CI/CM/CR meteorites, and was classified as an anomalous C2<sup>35</sup>.

The magnesite ( $\text{MgCO}_3$ ) mineral standard is from Brumado Bahia, Brazil. It was measured using a bi-directional VNIR spectrometer from 0.3  $\mu\text{m}$  to 2.5  $\mu\text{m}$  under ambient conditions, and a Nicolet Fourier transform infrared (FTIR) spectrometer in a biconical configuration from 1  $\mu\text{m}$  to 50  $\mu\text{m}$ , in an environment purged of  $\text{H}_2\text{O}$  and  $\text{CO}_2$  for 10–12 h, at Brown University's RELAB<sup>36</sup>. Spectra of dolomite, calcite and siderite were measured both at the University of Winnipeg PSF with an ASD Field Spec Pro HR spectrometer between 0.35  $\mu\text{m}$  and 2.5  $\mu\text{m}$ , and at Brown University with the RELAB facility between 2.0  $\mu\text{m}$  and 25  $\mu\text{m}$  (ref. 37).

Reflectance spectra of  $\text{NH}_4$ -montorillonite were measured using a Nicolet 740 FTIR spectrometer in a  $\text{H}_2\text{O}$ - and  $\text{CO}_2$ -purged environment. For further details see ref. 38. Spectra of  $\text{NH}_4$ -annite, cronstedtite and brucite were obtained through the RELAB online database.

Spectral fits including water ice and amorphous carbon were modelled using optical constants as derived in refs 39–43.

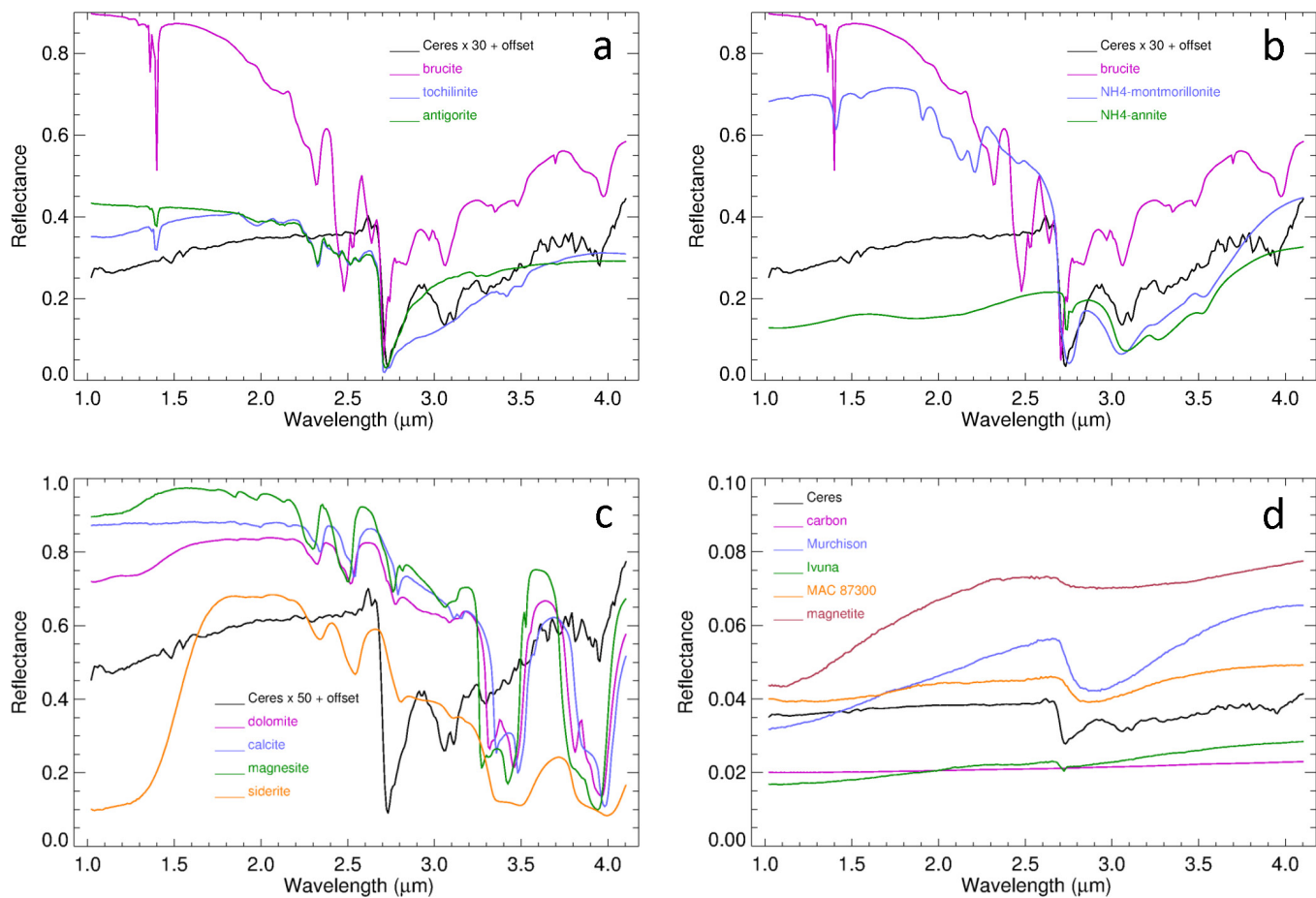
Spectra of magnetite were measured both at the University of Winnipeg PSF with an ASD Field Spec Pro HR spectrometer between 0.35  $\mu\text{m}$  and 2.5  $\mu\text{m}$ , and at Brown University with the RELAB facility between 2.0  $\mu\text{m}$  and 25  $\mu\text{m}$  (ref. 44).

Laboratory spectra of antigorite were measured after the samples were heated in a vacuum cell at different temperatures at the RELAB facility<sup>45</sup>.

**Sample size.** No statistical methods were used to predetermine sample size.

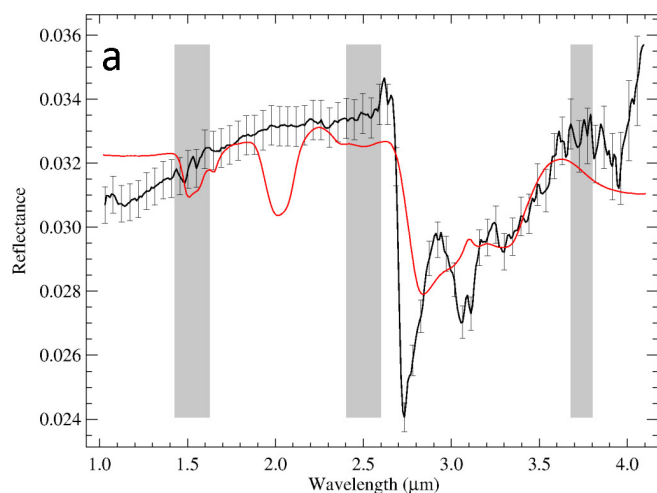
**Code availability.** We have opted not to make the code available because it is based on a well-known light scattering theory as described in Methods.

31. Carli, C., Ciarniello, M., Capaccioni, F., Serventi, G. & Sgavetti, M. Spectral variability of plagioclase-mafic mixtures (2): investigation of the optical constant and retrieved mineral abundance dependence on particle size distribution. *Icarus* **235**, 207–219 (2014).
32. Raponi, A. *Spectrophotometric Analysis of Cometary Nuclei from in situ Observations*. PhD thesis, TorVergata Univ. (2015).
33. Hiroi, T., Zolensky, M. E. & Pieters, C. M. Characterization of unusual CI/CM/CR meteorites from reflectance spectroscopy. *Lunar Planet. Sci. Conf.* **28**, 577 (1997).
34. Hiroi, T., Pieters, C. M., Zolensky, M. E. & Prinz, M. Reflectance spectra (UV–3  $\mu\text{m}$ ) of heated Ivuna (CI) meteorite and newly identified thermally metamorphosed CM chondrites. *Lunar Planet. Sci.* **27**, 551–552 (1996).
35. Hiroi, T. *et al.* Thermal metamorphism of the C, G, B, and F asteroids seen from 0.7  $\mu\text{m}$ , 3  $\mu\text{m}$ , and UV absorption strengths in comparison with carbonaceous chondrites. *Meteorit. Planet. Sci.* **31**, 321–327 (1996).
36. Bishop, J. L. *et al.* Coordinated spectral and XRD analyses of magnesite-nontronite-forsterite mixtures and implications for carbonates on Mars. *J. Geophys. Res. Planets* **118**, 635–650 (2013).
37. Cloutis, E. *et al.* Spectral reflectance properties of carbonates from terrestrial analogue environments: implications for Mars. *Planet. Space Sci.* **58**, 522–537 (2010).
38. Bishop, J. L. *et al.* Detection of soluble and fixed  $\text{NH}_4^+$  in clay minerals by DTA and IR reflectance spectroscopy: a potential tool for planetary surface exploration. *Planet. Space Sci.* **50**, 11–19 (2002).
39. Zubko, V. G. *et al.* Optical constants of cosmic carbon analogue grains — I. Simulation of clustering by a modified continuous distribution of ellipsoids. *Mon. Not. R. Astron. Soc.* **282**, 1321–1329 (1996).
40. Mastrapa, R. M. *et al.* Optical constants of amorphous and crystalline  $\text{H}_2\text{O}$ -ice in the near infrared from 1.1 to 2.6  $\mu\text{m}$ . *Icarus* **197**, 307–320 (2008).
41. Mastrapa, R. M. *et al.* Optical constants of amorphous and crystalline  $\text{H}_2\text{O}$ -ice: 2.5–22  $\mu\text{m}$  (4000–455  $\text{cm}^{-1}$ ) optical constants of  $\text{H}_2\text{O}$ -ice. *Astrophys. J.* **701**, 1347–1356 (2009).
42. Clark, R. N. *et al.* The surface composition of Iapetus: mapping results from Cassini VIMS. *Icarus* **218**, 831–860 (2012).
43. Warren, S. G. Optical constants of ice from the ultraviolet to the microwave. *Appl. Opt.* **23**, 1206–1225 (1984).
44. Cloutis, E. *et al.* Spectral reflectance properties of carbonaceous chondrites: 1. CI chondrites. *Icarus* **212**, 180–209 (2011).
45. Hiroi, T. K. & Zolensky, M. E. UV-Vis-NIR absorption features of heated phyllosilicates as remote-sensing clues of thermal histories of primitive asteroids. *Antarct. Meteorite Res.* **12**, 108–116 (1999).

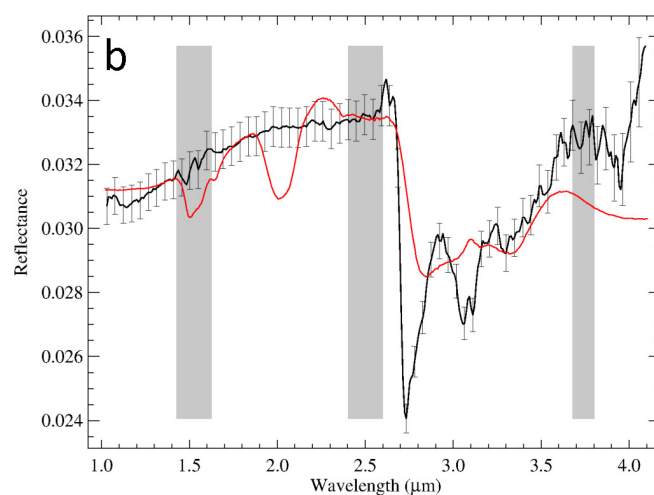


**Extended Data Figure 1 | End-member spectra used for spectral fit modelling.** **a**, Spectra of minerals showing a prominent absorption at  $2.7\ \mu\text{m}$ . **b**, Spectra of minerals showing a prominent  $3.06\text{--}3.1\ \mu\text{m}$

band. **c**, Spectra of carbonates; and **d**, spectra of dark components and carbonaceous chondrites used in the modelling. All panels show the Ceres spectrum (black) for comparison.

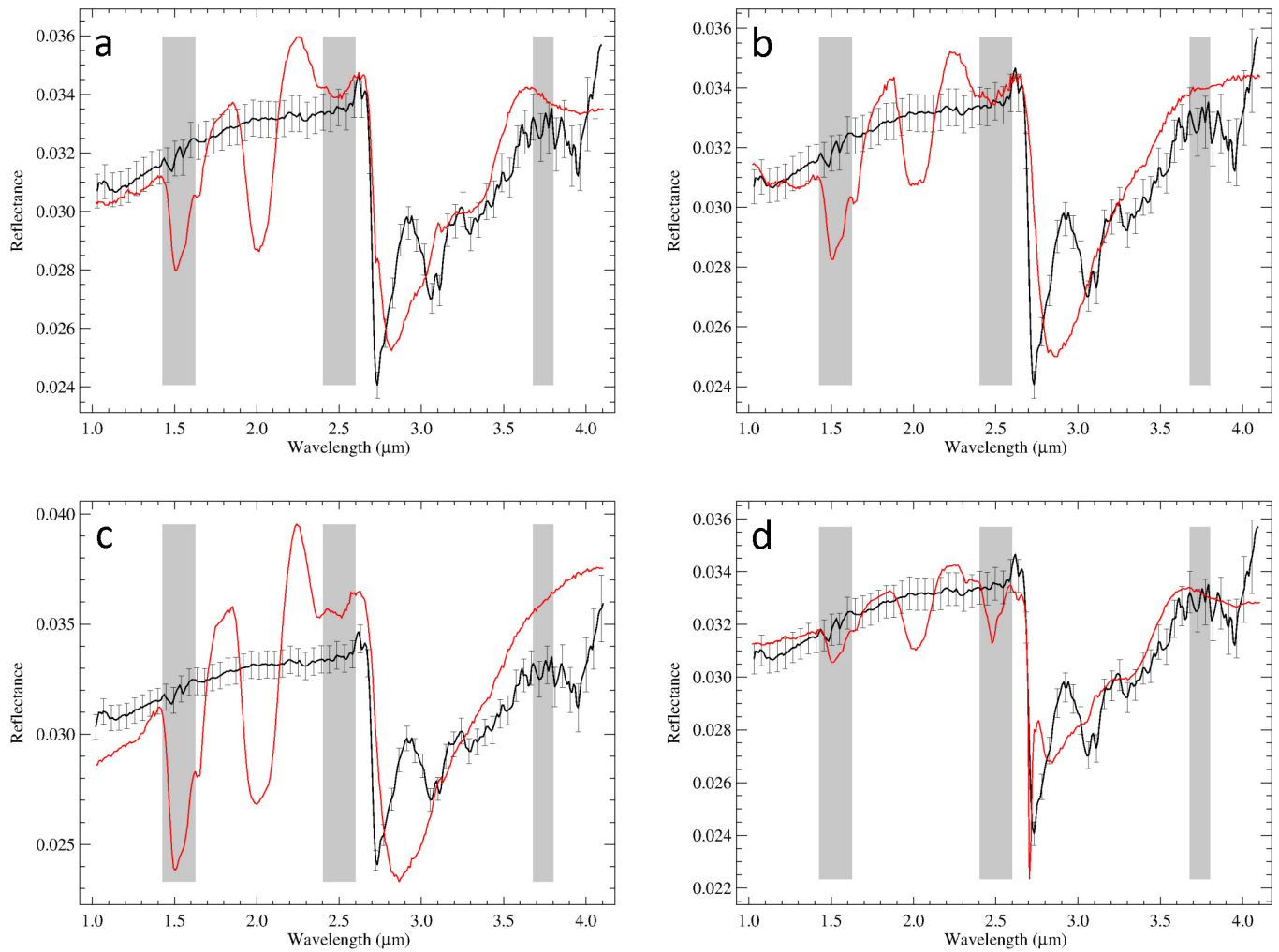


**Extended Data Figure 2 | Spectral fit of the spectrum of Ceres including water. a, b,** Red curves, results of the spectral fitting model using water ice and amorphous carbon (a) or water ice and magnetite (b). Ceres' spectrum



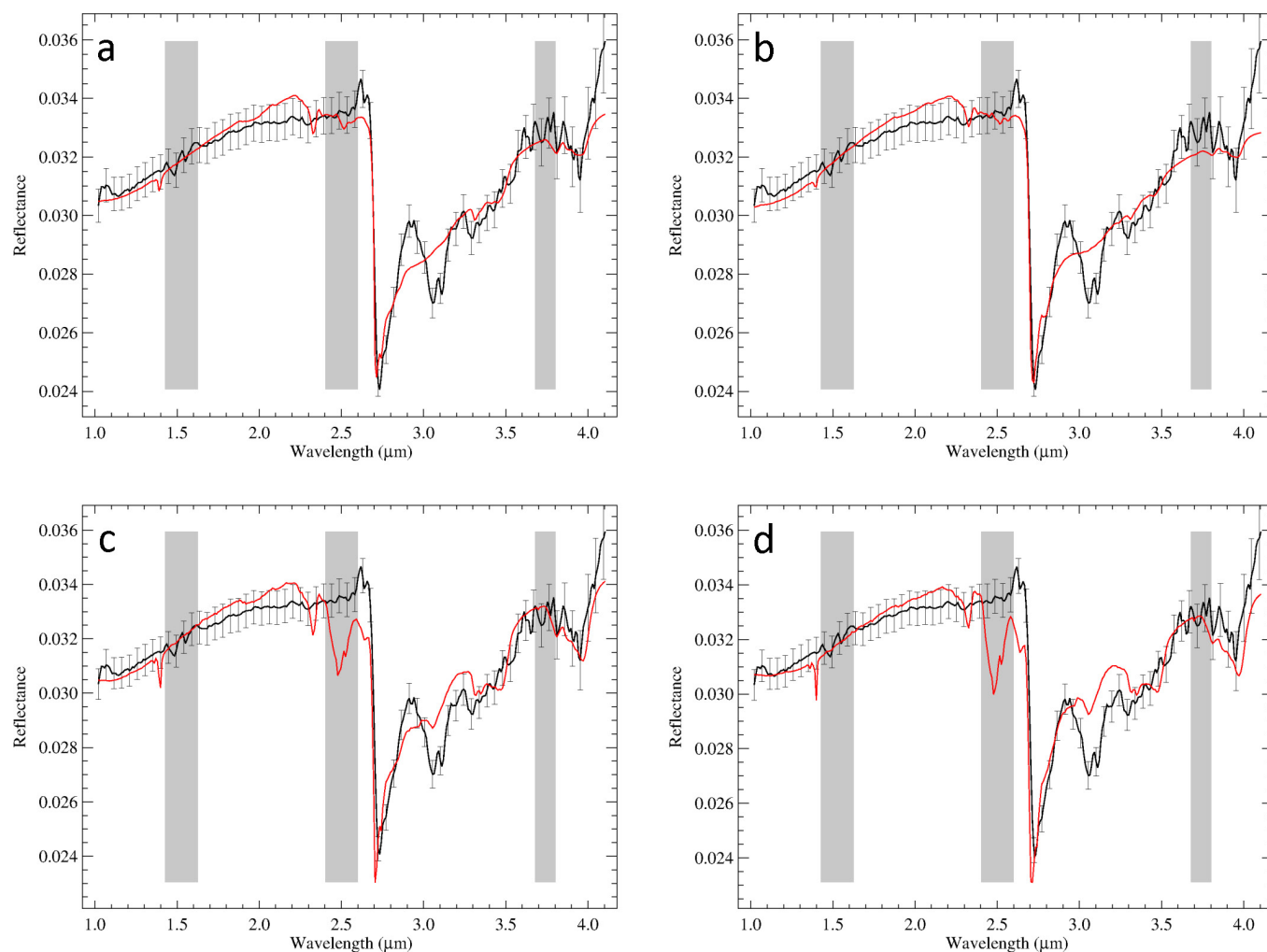
is in black. The error bars are calculated taking into account a mean absolute deviation of the calibration uncertainties along the 256 samples. Vertical grey bars represent regions of the detector affected by the filter junctions.





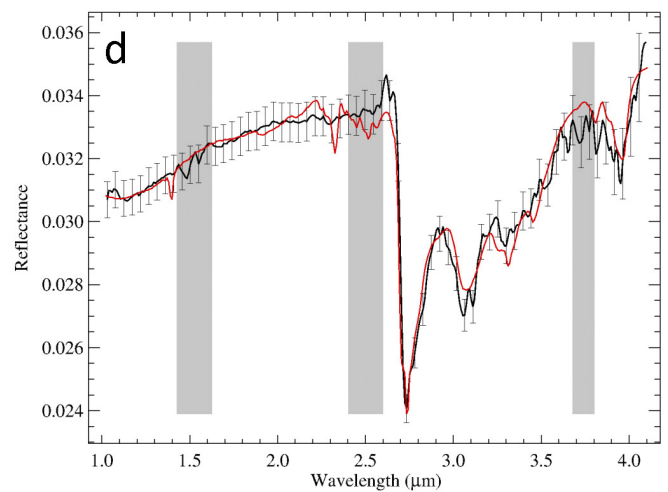
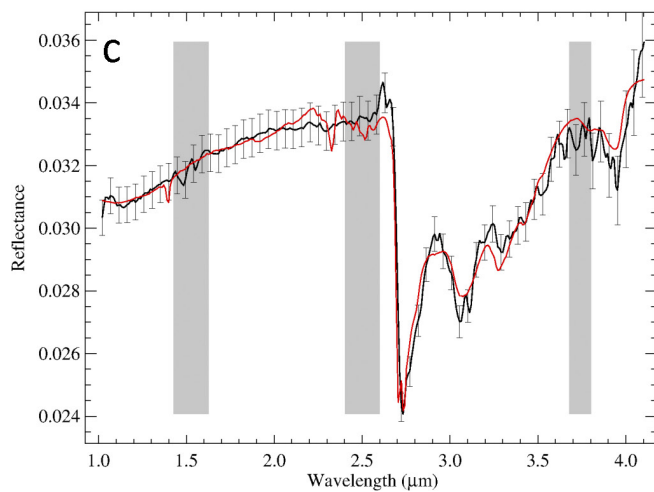
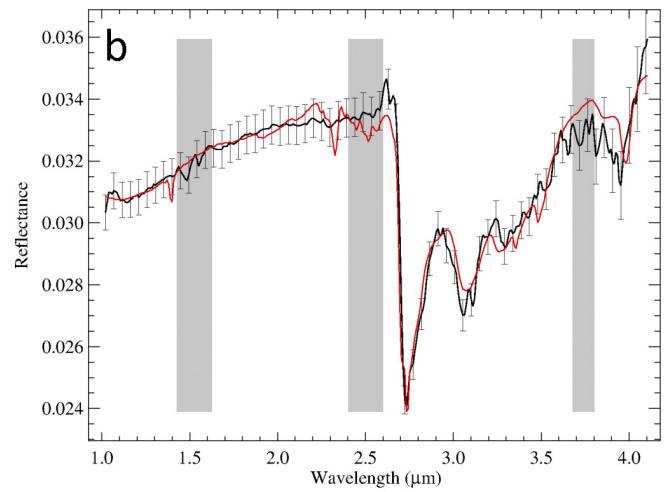
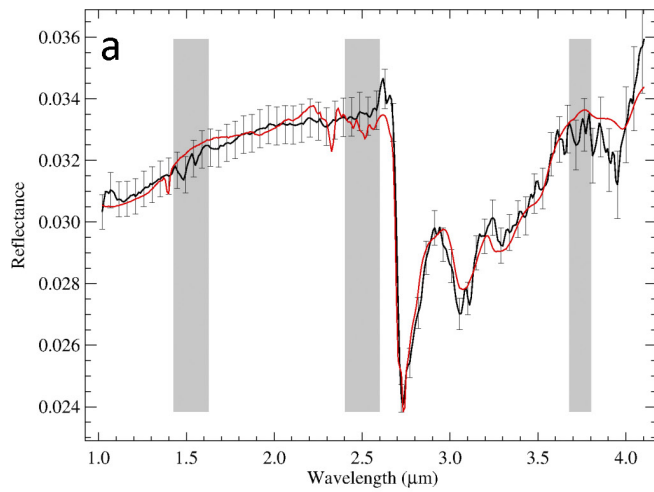
**Extended Data Figure 3 | Spectral fit of the spectrum of Ceres including carbonaceous chondrites. a–d,** Results of the spectral fitting model (red curves) using water ice and Ivuna CI chondrite (a; heated at 500 °C); water ice and MAC 87300 CM chondrite (b); water ice and Murchison CM

chondrite (c; heated at 500 °C); and water ice, Ivuna CI chondrite (heated at 500 °C) and brucite (d). Ceres' spectrum is in black. Error bars and vertical grey bars as in Extended Data Fig. 2.



**Extended Data Figure 4 | Spectral fit of the spectrum of Ceres including cronstedtite, tochilinite and brucite. a–d**, Results of the spectral fitting model (red curves) using tochilinite, cronstedtite, dolomite and magnetite (a); antigorite, cronstedtite, dolomite and magnetite (b); tochilinite, brucite,

dolomite and magnetite (c); and antigorite, brucite, dolomite and magnetite (d). Ceres's spectrum is in black. Error bars and vertical grey bars as in Extended Data Fig. 2.



**Extended Data Figure 5 | Spectral fit of the spectrum of Ceres including different carbonates. a–d,** Results of the spectral fitting model (red curves) using NH<sub>4</sub>-annite, antigorite, magnetite and siderite (a); NH<sub>4</sub>-annite, antigorite, magnetite and calcite (b); NH<sub>4</sub>-annite, antigorite,

magnetite and magnesite (c); and NH<sub>4</sub>-annite, antigorite, magnetite and dolomite (d). Ceres' spectrum is in black. Error bars and vertical grey bars as in Extended Data Fig. 2.



Extended Data Table 1 | End-members used in the mixtures

Mineral / Rock	Sample name	PI
amorphous carbon	/	Optical constant from [39]
Murchison	MB-TXH-064-HB	Takahiro Hiroi, Brown University Dept. of Geological Sciences
Ivuna (heated 500 °C)	MP-TXH-018	Takahiro Hiroi, Brown University, Dept. of Geological Sciences
MAC87300	MP-TXH-045	Takahiro Hiroi, Brown University, Dept. of Geological Sciences
magnetite	MG-EAC-002	Edward A. Cloutis, University of Winnipeg, Department of Geography
brucite	JB-JLB-944	Janice L. Bishop, SETI Institute
tochilinite	TO-TXH-001	Edward A. Cloutis, University of Winnipeg, Department of Geography
cronstedtite	CR-EAC-021	Edward A. Cloutis, University of Winnipeg, Department of Geography
antigorite (heated 500 °C)	AT-TXH-007	Takahiro Hiroi, Brown University, Dept. of Geological Sciences
NH4-montmorillonite	JB-JLB-189	Janice L. Bishop, SETI Institute
NH4-annite	AA-A1S-002	Alexander Smirnov, State University of New York, Dept. of Geosciences
dolomite	CB-EAC-003	Edward A. Cloutis, University of Winnipeg, Department of Geography
calcite	CA-EAC-010	Edward A. Cloutis, University of Winnipeg, Department of Geography
magnesite	JB-JLB-946	Janice L. Bishop, SETI Institute
siderite	CB-EAC-008	Edward A. Cloutis, University of Winnipeg, Department of Geography
water ice	/	Optical constant from [40,41,42,43]

All the spectra are from the RELAB database and are described in the Methods section. Note optical constants of amorphous carbon are from ref. 39, and those of water ice are from refs 40–43.

Extended Data Table 2 | Details of results obtained using different end-members

Endmembers	Abundance (%)	Grain size ( $\mu\text{m}$ )	$\chi^2$	figure
brucite	13	30	0.84 (a)	Fig. 3-a-b main text
cronstedtite	5	40	(2.85 – 4.1 $\mu\text{m}$ )	
Mg-carbonates	1	10	2.57 (b)	
dark component	83	*	(1.0 – 4.1 $\mu\text{m}$ )	
NH4 montmorillonite	6	**	0.70	Fig. 4-a main text
antigorite	5	90		
Mg-carbonates	5	10		
dark component	84	80		
NH4 annite	6	**	0.79	Fig. 4-b main text
antigorite	9	180		
Mg-carbonates	25	2		
dark component	60	80		
water ice	3	50	6.38	Fig. 2-a extended data
amorphous carbon	97	*	6.73	Fig. 2-b extended data
water ice	4	60		
magnetite	96	100	7.32	Fig. 3-a extended data
water ice	6	90	7.70	Fig. 3-b extended data
Ivuna CC (500 C)	94	*		
water ice	4	180	20	Fig. 3-c extended data
MAC CC	96	*		
water ice	10	300	3.30	Fig. 3-d extended data
Murchison CC	90	*		
water ice	4	60	1.18	Fig. 4-a extended data
Ivuna CC (500 C)	77	*		
brucite	19	10		
tochilinite	9	30		
cronstedtite	5	70	1.28	Fig. 4-b extended data
dolomite	1	70		
magnetite	85	180		
antigorite	10	50		
cronstedtite	7	30	2.14	Fig. 3-c extended data
dolomite	1	20		
magnetite	82	100		
tochilinite	6	90		
brucite	5	50	2.74	Fig. 4-d extended data
dolomite	2	100		
magnetite	87	80		
antigorite	6	100		
brucite	5	100	0.80	Fig. 5-a extended data
dolomite	2	100		
magnetite	87	80		
NH4 annite	7	**		
antigorite	10	160	0.82	Fig. 5-b extended data
magnetite	78	100		
siderite	5	10		
NH4 annite	7	**		
antigorite	10	190	0.79	Fig. 5-c extended data
magnetite	78	85		
calcite	5	15		
NH4 annite	7	**		
antigorite	11	190	0.79	Fig. 5-d extended data
magnetite	69	80		
magnesite	13	10		
NH4 annite	6	**		
antigorite	9	180	0.79	Fig. 5-d extended data
magnetite	60	80		
dolomite	25	2		

Spectra are given in Figs 3 and 4, and in Extended Data Figs 2–5.

\*Grain size cannot be determined unequivocally. In the case of very dark materials, albedo is weakly linked to the grain size.

\*\*The lack of grain size information for the measured reflectance spectra used as input prevents the determination of the effect of grain size on albedo.

# Nanoscale intimacy in bifunctional catalysts for selective conversion of hydrocarbons

Jovana Zečević<sup>1</sup>, Gina Vanbutsele<sup>2</sup>, Krijn P. de Jong<sup>1</sup> & Johan A. Martens<sup>2</sup>

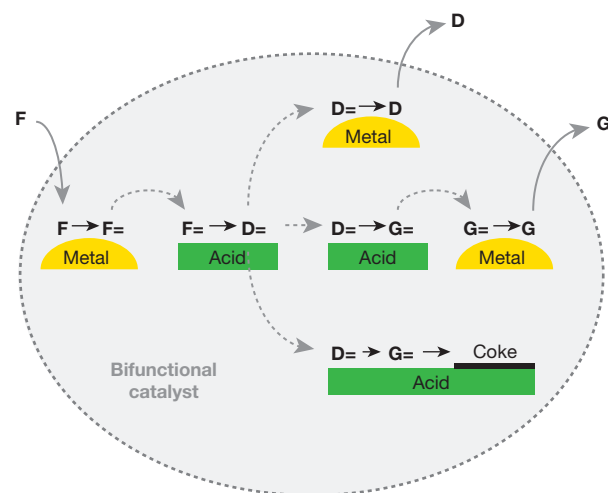
The ability to control nanoscale features precisely is increasingly being exploited to develop and improve monofunctional catalysts<sup>1–4</sup>. Striking effects might also be expected in the case of bifunctional catalysts, which are important in the hydrocracking of fossil and renewable hydrocarbon sources to provide high-quality diesel fuel<sup>5–7</sup>. Such bifunctional hydrocracking catalysts contain metal sites and acid sites, and for more than 50 years the so-called intimacy criterion<sup>8</sup> has dictated the maximum distance between the two types of site, beyond which catalytic activity decreases. A lack of synthesis and material-characterization methods with nanometre precision has long prevented in-depth exploration of the intimacy criterion, which has often been interpreted simply as ‘the closer the better’ for positioning metal and acid sites<sup>8–11</sup>. Here we show for a bifunctional catalyst—comprising an intimate mixture of zeolite Y and alumina binder, and with platinum metal controllably deposited on either the zeolite or the binder—that closest proximity between metal and zeolite acid sites can be detrimental. Specifically, the selectivity when cracking large hydrocarbon feedstock molecules for high-quality diesel production is optimized with the catalyst that contains platinum on the binder, that is, with a nanoscale rather than closest intimacy of the metal and acid sites. Thus, cracking of the large and complex hydrocarbon molecules that are typically derived from alternative sources, such as gas-to-liquid technology, vegetable oil or algal oil<sup>6,7</sup>, should benefit especially from bifunctional catalysts that avoid locating platinum on the zeolite (the traditionally assumed optimal location). More generally, we anticipate that the ability demonstrated here to spatially organize different active sites at the nanoscale will benefit the further development and optimization of the emerging generation of multifunctional catalysts<sup>12–15</sup>.

Bifunctional catalysts for the conversion of hydrocarbons contain metal sites next to acid sites. In the case of the conversion of normal alkanes, the metal sites catalyse (de)hydrogenation reactions; the acid sites catalyse isomerization and cracking reactions, to obtain iso-alkanes of the same or reduced molecular weight, respectively<sup>16</sup>. The critical properties of these catalysts are the ratio of metal sites relative to acid sites, and the proximity of both sites to each other<sup>8,17</sup>. This latter aspect is further illustrated by the schematic presentation of hydrocracking reactions in Fig. 1. The normal alkene reaction intermediate has to diffuse from the metal site to the acid site; the iso-alkene reaction intermediate must diffuse the other way. If the distances are too large and diffusivity too low, concentration gradients develop and catalytic activity decreases. The intimacy criterion put forward by Weisz<sup>8</sup> in 1962 provides the maximum distance between both types of site, beyond which catalytic activity decreases. Since then it has been realized that, even more importantly, selectivity might also be affected by intimacy<sup>9–11,18</sup>: if the distances between metal and acid sites are large, iso-alkenes might undergo secondary reactions, giving rise to gas and to coke. Systematic quantitative studies on intimacy effects have been carried out only at the macroscale<sup>8,11,18</sup> (tens of micrometres and larger); meanwhile, in many other studies the metal function has

been put as close as possible to the acid sites<sup>6,9,19</sup> (in particular in the case of metal dispersed in acid zeolite). These findings have led to the general belief that, as far as the intimacy of bifunctional catalysts goes, ‘the closer the better’.

Here we report a new methodology for controlling and imaging the intimacy of metal and acid sites at the nanoscale, with large positive effects on selectivity. Our strategy for controlling the distance between metal and acid sites starts with an intimate mixture of 50 wt% mesoporous zeolite Y (designated Y; this provides the acid sites) and 50 wt%  $\gamma$ -alumina binder (designated as A), obtained by extrusion and referred to as Y/A. Energy dispersive X-ray (EDX) spectroscopy maps of a 70-nm-thick ultramicrotomed slice of Y/A (Fig. 2) show the two components mixed at the nanoscale, with zeolite crystals (of diameter 500–1,000 nm) being surrounded by at least one layer of alumina, and alumina regions not exceeding 5  $\mu$ m in length. Porosity, apparent as black voids in the EDX map (Fig. 2a) and as revealed by nitrogen physisorption (Extended Data Fig. 1), enables molecular transport via diffusion throughout the extrudates.

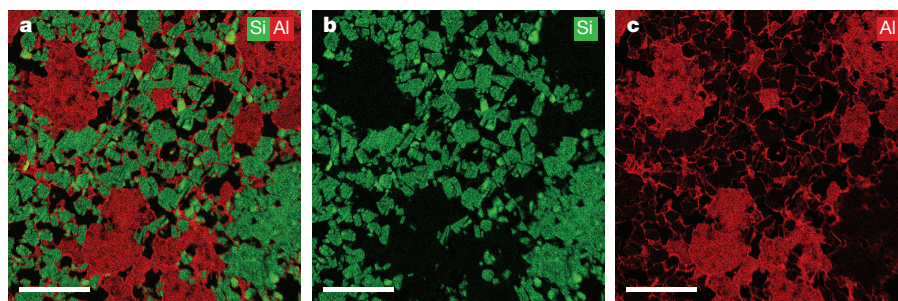
We then exploit the difference between ion exchange and electrostatic adsorption<sup>20,21</sup> to direct the subsequent deposition of platinum (Pt) metal precursor: ion exchange at pH5 with an aqueous solution of  $\text{Pt}(\text{NH}_3)_4(\text{NO}_3)_2$  results in Pt cations being located exclusively on the zeolite component, whereas electrostatic adsorption at pH3 from



**Figure 1 | Scheme of hydrocracking reactions that use a bifunctional catalyst.** Feed normal alkane molecules (F) are dehydrogenated on a metal surface, producing alkene intermediates (F=). The alkenes diffuse to zeolite Brønsted acid sites, on which they undergo acid-catalysed skeletal isomerization, which can be followed by one (D=) or more (G=) cracking events, sometimes leading to coke formation. Isomerized (D=) and/or cracked (D=, G=) alkene intermediates diffuse to the metal site and are hydrogenated to form isomerized or cracked products D (diesel) and G (gas).

<sup>1</sup>Inorganic Chemistry and Catalysis, Debye Institute for Nanomaterials Science, Utrecht University, Universiteitsweg 99, 3584 CG Utrecht, The Netherlands. <sup>2</sup>Centre for Surface Chemistry and Catalysis, KU Leuven, Celestijnenlaan 200F Postbus 2461, B-3001 Leuven, Belgium.





**Figure 2 | Distribution of zeolite Y and alumina components within Y/A extrudates.** **a**, EDX map of a 70-nm-thick section of a Y/A extrudate, showing the location of silicon (Si, green) and aluminium (Al, red), which are indicative of the presence of the zeolite Y and alumina components,

respectively, of Y/A. **b**, EDX map showing only the Si (green) signal, corresponding to the presence of zeolite Y. **c**, EDX map showing only the Al (red) signal, most pronounced in regions containing pure alumina phase. Scale bars, 2  $\mu\text{m}$ .

a solution of  $\text{H}_2\text{PtCl}_6$  deposits Pt anions exclusively on the alumina component (see Methods). The resulting samples, Pt-Y/A and Pt-A/Y, are then either heat-treated in  $\text{H}_2$ , or heat-treated in  $\text{H}_2$  followed by 1 vol%  $\text{O}_2$  in  $\text{N}_2$ , respectively, to generate Pt metal particles with a size of about 3 nm (refs 22, 23). Inductively coupled plasma (ICP) analysis shows the Pt-Y/A and Pt-A/Y samples to have comparable Pt loading of 0.6 wt% and 0.7 wt%, respectively, which suffices to maintain the metal–acid balance needed for so-called ideal hydrocracking<sup>11</sup>.

High-angle annular dark-field scanning transmission electron microscopy (HAADF-STEM) imaging and EDX elemental mapping of 70-nm-thick ultramicrotomed sections of Pt-Y/A (Fig. 3a, b and Extended Data Fig. 2a) confirms that Pt particles with a narrow size distribution, around 2.5 nm, are exclusively present in the zeolite crystals. Cutting the extrudates with ultramicrotomy to produce 70-nm-thick slices was crucial, to expose regions that contain purely the zeolite or alumina components. The images in Fig. 3a, b and electron tomography observations of Pt-Y/A (Extended Data Fig. 3a) reveal Pt nanoparticles

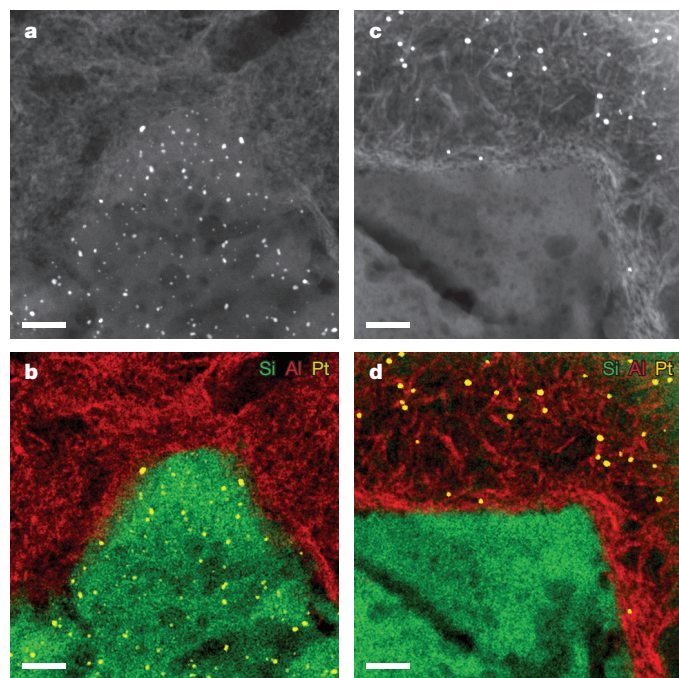
located within the zeolite crystals even though they exceed the size of the micropores (roughly 1 nm), as reported previously<sup>22</sup>. In the Pt-A/Y sample, Pt particles with an average size of 3.5 nm and narrow size distribution reside exclusively on the alumina phase of the extrudates (Fig. 3c, d and Extended Data Fig. 2b). The absence of Pt particles in the zeolite crystals was confirmed by electron tomography (Extended Data Fig. 3b).

The two catalysts thus have well defined structures, with Pt particles of comparable size located either inside the zeolite pores in closest proximity to the zeolite acid sites (in the Pt-Y/A sample), or within a nanoscale distance from the acid sites of the zeolite when located on the binder (in Pt-A/Y). Given that only about 0.6 wt% of Pt was loaded on the pre-shaped Y/A extrudates, no change in textural properties is expected compared with pristine Y/A extrudates. Temperature-programmed desorption of ammonia also indicates very similar acidities for the two catalysts (Extended Data Fig. 4). The two catalyst samples, with all structural parameters being identical except for the Pt location, therefore allow us to evaluate the impact of nanoscale intimacy on hydroconversion activity and selectivity.

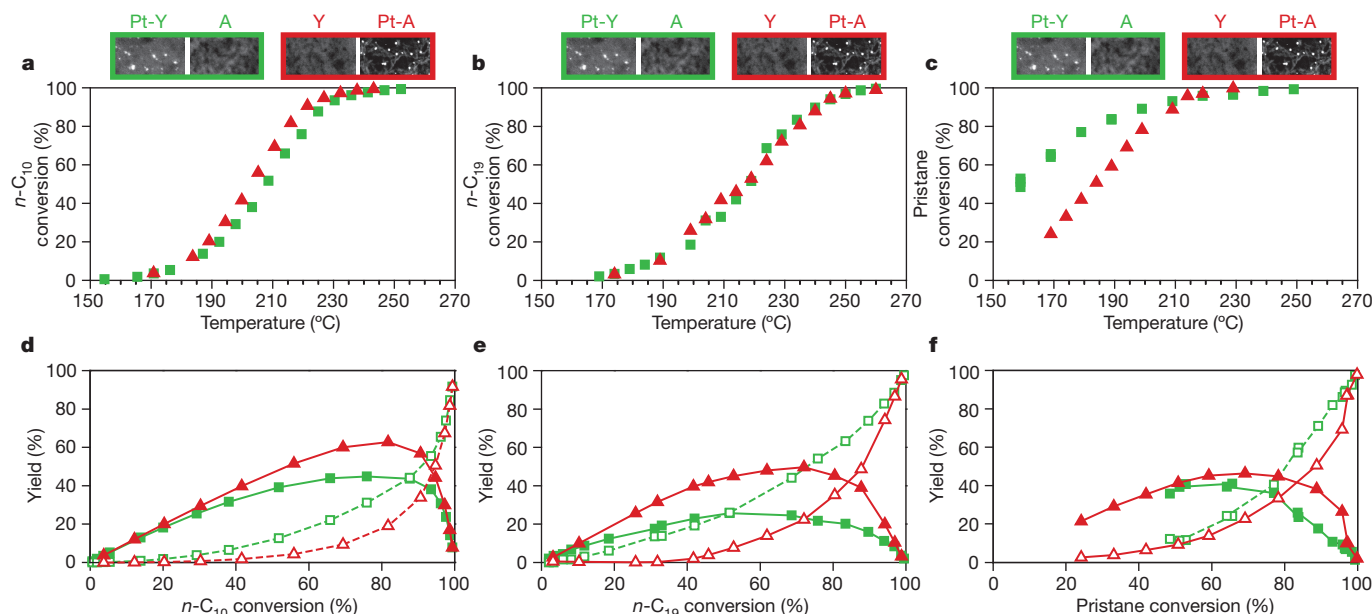
The model feedstocks in our catalysis experiments are *n*-decane (*n*-C<sub>10</sub>), *n*-nonadecane (*n*-C<sub>19</sub>) and pristane (2,6,10,14-tetramethylpentadecane, *i*-C<sub>19</sub>), with the two catalysts showing very similar *n*-C<sub>10</sub> and *n*-C<sub>19</sub> conversion activity (Fig. 4a, b). But Pt-Y/A, with its metal and zeolite acid sites in closest proximity, exhibits improved activity over Pt-A/Y in the conversion of multibranched *i*-C<sub>19</sub> feed molecules (Fig. 4c). Assuming a typical, normal alkene diffusivity in zeolite Y of  $1.0 \times 10^{-9} \text{ m}^2 \text{ s}^{-1}$ , the maximum intersite distance for *n*-alkane conversion has been estimated to be 4.5  $\mu\text{m}$ , using the formula for the intimacy criterion<sup>24</sup>. With Pt on the alumina binder, the maximum distance between sites of about 500 nm is below this estimate and in line with the identical catalytic activities of the two catalysts for *n*-C<sub>10</sub> and *n*-C<sub>19</sub> conversion (Fig. 4a, b). For *i*-C<sub>19</sub>, however, the lower activity of Pt-A/Y points to its activity being limited by transport of the bulky tetramethylbranched *i*-C<sub>19</sub> alkene molecules, with their estimated diffusivity of only  $1.0 \times 10^{-11} \text{ m}^2 \text{ s}^{-1}$ .

In contrast to their conversion activities, the selectivities of the two catalysts differ remarkably with all three feeds (Fig. 4d–f): the Pt-A/Y catalyst always produces a much higher yield of skeletal isomers, thus illustrating the importance of nanoscale intimacy between (de)hydrogenation and acid functions for the selectivity of bifunctional catalysts. Skeletal branching is succeeded by cracking of the carbon chains. Undesired multiple cracking leads to C<sub>3</sub>–C<sub>4</sub> products<sup>19</sup>, indicated as gas (Fig. 1). The cracking products formed at the same cracking conversion rate were identical for the two catalysts when the feedstock was *n*-C<sub>10</sub>, but markedly different when the feedstock was *n*-C<sub>19</sub> (Extended Data Fig. 5). In case of *n*-C<sub>19</sub>, the product distribution was shifted towards the more desired range of middle distillates (C<sub>10</sub>–C<sub>19</sub>) when using the Pt-A/Y catalyst; the Pt-Y/A catalyst produced more gas (C<sub>3</sub>–C<sub>4</sub>) and naphtha (C<sub>5</sub>–C<sub>9</sub>) (Extended Data Fig. 5b).

High isomerization yield and limited secondary cracking are characteristics of ideal hydrocracking. In this regard, the Pt-A/Y catalyst, with



**Figure 3 | Controlled deposition of platinum (Pt) on either the zeolite Y or the alumina component of Y/A extrudates.** **a**, **b**, Pt-Y/A. The HAADF-STEM image (**a**) is of a 70-nm-thick section of a Pt-Y/A sample, with Pt particles of ~2.5 nm residing exclusively within the zeolite crystals, as evident from the EDX map (**b**) showing Pt (yellow), Si (green) and Al (red) signals. **c**, **d**, Pt-A/Y. Shown are an HAADF-STEM image (**c**) and EDX map (**d**) of a 70-nm-thick section of Pt-A/Y, with Pt particles (yellow) of ~3.5 nm residing exclusively on the alumina platelets (red) while the zeolite crystals (green) are empty. Scale bars, 50 nm.



**Figure 4 | Impact of nanoscale intimacy on hydrocracking activity and selectivity.** **a–c**, The percentage of alkane feedstock converted is shown against the reaction temperature for Pt-Y/A (green squares) or Pt-A/Y (red triangles) catalysts. The feedstocks are: **a**, *n*-decane (*n*-C<sub>10</sub>); **b**, *n*-nonadecane (*n*-C<sub>19</sub>); and **c**, pristane (*i*-C<sub>19</sub>). **d–f**, Product yields from **d**, *n*-decane, **e**, *n*-nonadecane, and **f**, pristane feedstocks, when using

Pt-Y/A (green) or Pt-A/Y (red) catalysts. Solid lines and filled symbols show the yields of isomerized products; dashed lines and open symbols show cracked products. Experiments were performed at a pressure of 0.45 MPa and a H<sub>2</sub>/hydrocarbon molar ratio of 214 (*n*-decane); or a pressure of 0.65 MPa and a H<sub>2</sub>/hydrocarbon molar ratio of 14.6 (*n*-nonadecane and pristane).

the metal function residing on alumina and thus more distant from the zeolite acid sites, outperforms the Pt-Y/A catalyst, in which Pt nanoparticles are located inside the zeolite micropores and are thus in more intimate contact with the acid sites. These findings contradict previous suggestions that having the metal and acid functions as close as possible promotes immediate hydrogenation of alkene intermediates and prevents undesired consecutive reactions such as multiple cracking<sup>11</sup> (Fig. 1). We note that in the earlier work *n*-hexadecane was often used as the feedstock, whereas we used a range of hydrocarbons (*n*-C<sub>10</sub>, *n*-C<sub>19</sub>, *i*-C<sub>19</sub>), allowing us to delineate the effects of intimacy on activity and selectivity. Moreover, in earlier studies a systematic variation in intimacy was achieved by mixing separately functionalized powders at the macroscale (tens of micrometres and more)<sup>8,11,18</sup>, whereas we used synthesis and characterization methods that allow for direct control of intimacy at the nanoscale and for observation of the reported large impact.

When using the catalyst Pt-Y/A, where the Pt metal function resides inside the zeolite crystals and in the highest possible intimacy with acid sites, feed molecules have to diffuse through the micropores to reach the metal sites. Once alkene intermediates are formed on the metal, we envisage them remaining trapped in the zeolite micropores because of their strong adsorption on acid sites and concomitant slow diffusion and longer residence times. This increases the probability of secondary reactions, particularly multiple cracking. With the catalyst Pt-A/Y, in which metal nanoparticles are located outside the zeolite crystals on the more easily accessible alumina binder, alkene intermediates formed on the metal sites diffuse through wide pores to the zeolite, where they immediately undergo isomerization. We suggest that alkene intermediates react mainly in the outer zeolite layers near the crystal surface, from where they can rapidly diffuse back to the metal sites on the alumina. This mechanism has been proposed for medium-pore zeolites exhibiting so-called pore mouth catalysis<sup>25</sup>, and is now extended to reactions of heavy alkanes on large-pore zeolite Y.

Hydrocracking catalysts that involve sulfided nickel molybdenum on alumina–zeolite, in which metal (sulfide) particles are mainly on the alumina binder<sup>9,26</sup>, are widely used and probably already meet the requirement that the optimal location for the metal function of

bifunctional catalysts is not in the micropores of the zeolite, as long perceived, but rather on the binder or on the surface/in the mesopores of the zeolite. But with the worldwide growing interest in alternative feedstocks, such as those based on gas-to-liquids<sup>6</sup> technology and on vegetable oil and algal oil<sup>7</sup>, molecules of even larger complexity and size present new challenges to the process of hydroconversion for high-quality diesel production. Improving hydrocracking catalysts generally, and in such challenging applications, will undoubtedly benefit from full control over the size and location of the metal function in complex catalysts, complementing recent efforts that have focused on nanozeolites<sup>10,27</sup>, hierarchical zeolites<sup>28</sup>, metal particle size<sup>29</sup> and nano-scale characterization<sup>30</sup> to improve our understanding. Our control strategy—and the methodology used to achieve it—can be extended to other bifunctional catalysts<sup>12–15</sup>, and provides the opportunity for tailoring catalyst properties in accordance with the specific requirements of the target process, beyond a simple ‘the closer the better’ notion of intimacy.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 6 April; accepted 19 October 2015.**

- Torres Galvis, H. M. *et al.* Supported iron nanoparticles as catalysts for sustainable production of lower olefins. *Science* **335**, 835–838 (2012).
- Cui, C., Gan, L., Heggen, M., Rudi, S. & Strasser, P. Compositional segregation in shaped Pt alloy nanoparticles and their structural behaviour during electrocatalysis. *Nature Mater.* **12**, 765–771 (2013).
- Prieto, G., Zečević, J., Friedrich, H., de Jong, K. P. & de Jongh, P. E. Towards stable catalysts by controlling collective properties of supported metal nanoparticles. *Nature Mater.* **12**, 34–39 (2013).
- Luo, W. *et al.* High performing and stable supported nano-alloys for the catalytic hydrogenation of levulinic acid to  $\gamma$ -valerolactone. *Nature Commun.* **6**, 6540 (2015).
- Weitkamp, J. Catalytic hydrocracking—mechanisms and versatility of the process. *ChemCatChem* **4**, 292–306 (2012).
- Bouchy, C., Hastoy, G., Guillon, E. & Martens, J. A. Fischer-Tropsch waxes upgrading via hydrocracking and selective hydroisomerization. *Oil Gas Sci. Technol.* **64**, 91–112 (2009).
- Tran, N. H. *et al.* Catalytic upgrading of biorefinery oil from micro-algae. *Fuel* **89**, 265–274 (2010).
- Weisz, P. B. Polyfunctional heterogeneous catalysis. *Adv. Catal.* **13**, 137–190 (1962).

9. Francis, J. *et al.* Design of improved hydrocracking catalysts by increasing the proximity between acid and metallic sites. *Appl. Catal. A* **409–410**, 140–147 (2011).
10. Kim, J., Kim, W., Seo, Y., Kim, J.-C. & Ryoo, R. n-Heptane hydroisomerization over Pt/MFI zeolite nanosheets: effects of zeolite crystal thickness and platinum location. *J. Catal.* **301**, 187–197 (2013).
11. Batalha, N., Pinard, L., Bouchy, C., Guillon, E. & Guisnet, M. n-Hexadecane hydroisomerization over Pt-HBEA catalysts. Quantification and effect of the intimacy between metal and protonic sites. *J. Catal.* **307**, 122–131 (2013).
12. Huber, G. W., Chheda, J. N., Barrett, C. J. & Dumesic, J. A. Production of liquid alkanes by aqueous-phase processing of biomass-derived carbohydrates. *Science* **308**, 1446–1450 (2005).
13. Gao, J. *et al.* Identification of molybdenum oxide nanostructures on zeolites for natural gas conversion. *Science* **348**, 686–690 (2015).
14. Yamada, Y. *et al.* Nanocrystal bilayer for tandem catalysis. *Nature Chem.* **3**, 372–376 (2011).
15. Peng, X. *et al.* Impact of hydrogenolysis on the selectivity of the Fischer–Tropsch synthesis: diesel fuel production over mesoporous zeolite-Y-supported cobalt nanoparticles. *Angew. Chem. Int. Edn* **54**, 4553–4556 (2015).
16. Coonradt, H. L. & Garwood, W. E. Mechanism of hydrocracking. *Ind. Eng. Chem. Process Des. Dev.* **3**, 38–45 (1964).
17. Guisnet, M. “Ideal” bifunctional catalysis over Pt-acid zeolites. *Catal. Today* **218–219**, 123–134 (2013).
18. Zhang, A., Nakamura, I., Aimoto, K. & Fujimoto, K. Isomerization of n-pentane and other light hydrocarbons on hybrid catalyst. Effect of hydrogen spillover. *Ind. Eng. Chem. Res.* **34**, 1074–1080 (1995).
19. Martens, J. A., Jacobs, P. A. & Weitkamp, J. Attempts to rationalize the distribution of hydrocracked products. II. Relative rates of primary hydrocracking modes of long chain paraffins in open zeolites. *Appl. Catal.* **20**, 283–303 (1986).
20. Schreier, M., Teren, S., Belcher, L., Regalbuto, J. R. & Miller, J. T. The nature of ‘overexchanged’ copper and platinum on zeolites. *Nanotechnology* **16**, S582–S591 (2005).
21. Cho, H.-R. & Regalbuto, J. R. The rational synthesis of Pt-Pd bimetallic catalysts by electrostatic adsorption. *Catal. Today* **246**, 143–153 (2015).
22. Zečević, J., van der Eerden, A. M. J., Friedrich, H., de Jongh, P. E. & de Jong, K. P. Heterogeneities of the nanostructure of platinum/zeolite Y catalysts revealed by electron tomography. *ACS Nano* **7**, 3698–3705 (2013).
23. Boubnov, A. *et al.* Structure–activity relationships of Pt/Al<sub>2</sub>O<sub>3</sub> catalysts for CO and NO oxidation at diesel exhaust conditions. *Appl. Catal. B* **126**, 315–325 (2012).
24. Martens, J. A. & Jacobs, P. A. Introduction to acid catalysis with zeolites in hydrocarbon reactions. *Stud. Surf. Sci. Catal.* **137**, 633–671 (2001).
25. Martens, J. A. *et al.* Selective isomerization of hydrocarbon chains on external surfaces of zeolite crystals. *Angew. Chem. Int. Edn* **34**, 2528–2530 (1995).
26. Landau, M. V. *et al.* Hydrocracking of heavy vacuum gas oil with a Pt/H-beta-Al<sub>2</sub>O<sub>3</sub> catalyst: effect of zeolite crystal size in the nanoscale range. *Ind. Eng. Chem. Res.* **42**, 2773–2782 (2003).
27. Awala, H. *et al.* Template-free nanosized faujasite-type zeolites. *Nature Mater.* **14**, 447–451 (2015).
28. de Jong, K. P. *et al.* Zeolite Y crystals with trimodal porosity as ideal hydrocracking catalysts. *Angew. Chem. Int. Edn* **49**, 10074–10078 (2010).
29. Van Santen, R. A. Complementary structure sensitive and insensitive catalytic relationships. *Acc. Chem. Res.* **42**, 57–66 (2009).
30. Gommers, C. J. *et al.* Mesoscale characterization of nanoparticles distribution using X-ray scattering. *Angew. Chem. Int. Edn* **54**, 11804–11808 (2015).

**Acknowledgements** This work has been supported by the NRSC-C and the European Research Council, EU FP7 ERC Advanced Grant no. 338846. (J.Z. and K.P.deJ.), and by the Flemish government via the Methusalem program (J.A.M.). We thank J. D. Meeldijk for assistance with ultramicrotomy and electron microscopy; M. Rigutto (Shell Projects and Technology) for the zeolite–alumina extrudates; R. Oord and J. Ruiz-Martínez for ammonia TPD measurements; and M. De Prins and S. Radhakrishnan for help with the catalytic experiments.

**Author Contributions** J.Z. synthesized and characterized the samples, and drafted the manuscript. G.V. performed hydrocracking tests. K.P.deJ. and J.A.M. contributed to experimental design, data analysis and manuscript writing.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.A.M. ([johan.martens@biw.kuleuven.be](mailto:johan.martens@biw.kuleuven.be)) or K.P.deJ. ([k.p.dejong@uu.nl](mailto:k.p.dejong@uu.nl)).



## METHODS

**Catalyst synthesis.** Cylindrical extrudates with length  $\sim 8$  mm and diameter  $\sim 2$  mm, containing 50 wt% of mesoporous zeolite Y (Zeolyst CBV 760, Si/Al = 30 at/at) and 50 wt% of  $\gamma$ -alumina, were obtained from Shell Projects and Technology and are designated as Y/A. Extrudates were crushed and sieved to produce particles of 0.2–0.5 mm and, as such, were used as a support for Pt.  $\text{Pt}(\text{NH}_3)_4(\text{NO}_3)_2$  (99.995% purity) was purchased from Sigma-Aldrich, and  $\text{H}_2\text{PtCl}_6 \cdot 6\text{H}_2\text{O}$  (40 wt% Pt) was purchased from Merck.

**Introduction of Pt on the zeolite component—Pt-Y/A catalyst.** 1.485 g of 0.2–0.5-mm Y/A particles were suspended in 450 ml Milli-Q water at room temperature and stirred for 1 hour, after which the pH of the suspension was 6.4. Aqueous solution (50 ml) containing 29.8 mg of  $\text{Pt}(\text{NH}_3)_4(\text{NO}_3)_2$  was added dropwise to the suspension, after which the pH dropped to 5.3. At these conditions, ion exchange of zeolite protons with  $\text{Pt}(\text{NH}_3)_4^{2+}$  was dominant over electrostatic adsorption on alumina, because the latter surface at  $\text{pH} < 9$  is positively charged and does not interact strongly with Pt cations<sup>20</sup>. ICP analysis showed that, under these synthesis conditions, ion exchange led to Pt loading of 0.6 wt% (intake 1.0 wt%) with respect to Y/A. After addition of the Pt precursor, the suspension was stirred for another 3 hours, after which the pH dropped to 4.9. The suspension was filtered and washed with Milli-Q water, and dried in air overnight at 120 °C. Dried catalyst was reduced in a flow of  $\text{H}_2$  (gas hourly space velocity, GHSV,  $\sim 2,580 \text{ h}^{-1}$ ) for 3 h at 600 °C, using a ramp of  $5^\circ \text{C min}^{-1}$ .

**Introduction of Pt on the alumina component—Pt-A/Y catalyst.** A suspension containing 1.491 g of 0.2–0.5-mm Y/A extrudates particles in 450 ml Milli-Q water was stirred for 1 hour at room temperature, after which the pH of the suspension was lowered from 6.2 to 2.6 by adding 1 M HCl solution. To achieve the same Pt loading as for Pt-Y/A, we added 50 ml of aqueous solution containing 22.5 mg of  $\text{H}_2\text{PtCl}_6 \cdot 6\text{H}_2\text{O}$  salt ( $\text{pH} = 2.7$ ) dropwise to the suspension. At the applied pH range, it is expected that  $\text{PtCl}_6^{2-}$  will be adsorbed onto the positively charged alumina surface<sup>21</sup>, whereas no ion exchange or electrostatic adsorption on the zeolite will take place. ICP analysis confirmed that electrostatic adsorption was complete and that Pt loading was 0.7 wt% (intake 0.6 wt%) with respect to Y/A. After 3 hours of stirring ( $\text{pH}$  increased to 3.3), the suspension was filtered and washed with Milli-Q water, and dried in air overnight at 120 °C. Dried catalyst was reduced in a flow of  $\text{H}_2$  (GHSV  $\sim 2,580 \text{ h}^{-1}$ ) for 3 hours at 600 °C, with a heating ramp of  $5^\circ \text{C min}^{-1}$ , and then heat treated for 1 hour at 600 °C in a flow of 1 vol%  $\text{O}_2$  in  $\text{N}_2$  (GHSV  $\sim 2,580 \text{ h}^{-1}$ ).

When preparing smaller amounts of catalysts, we decreased the volume of Milli-Q water in the suspension proportionally, while a GHSV of  $\sim 3,300 \text{ h}^{-1}$  was used for heat treatments; however, we observed no impact of these changes on catalyst structure.

**Catalyst characterization.**  $\text{N}_2$  physisorption measurements were performed on a Micromeritics TriStar 3000 at liquid nitrogen temperature. The sample was dried for 14 h at 300 °C in  $\text{N}_2$  flow before measurement.

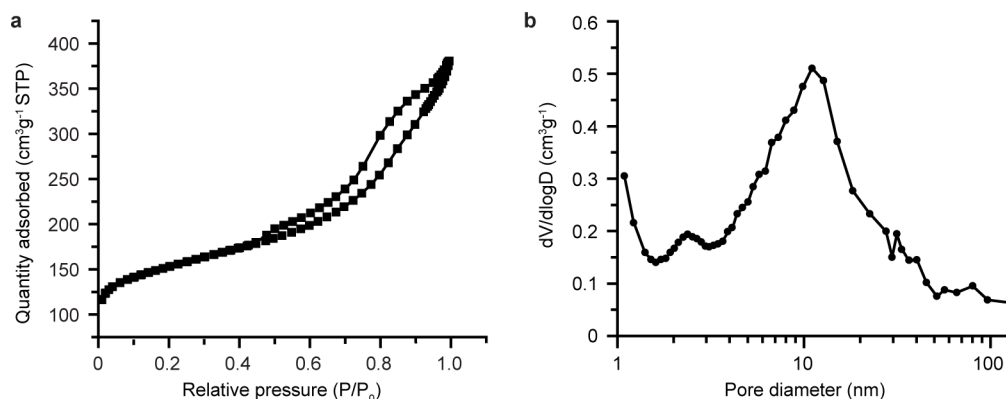
Temperature-programmed desorption (TPD) of ammonia was performed on a Micromeritics AutoChem II equipped with a thermal conductivity detector (TCD). Before TPD, 0.1 g of catalyst was dried in He for 1 hour at 600 °C, with a heating ramp of  $10^\circ \text{C min}^{-1}$ . The temperature was then decreased to 100 °C and ammonia (10 vol% in He) was adsorbed in a pulse-wise manner until oversaturated, after which the physisorbed ammonia was removed by flowing He for 1 h at 100 °C. The ammonia desorption was monitored until 600 °C with a ramp of  $5^\circ \text{C min}^{-1}$ .

ICP analysis of Pt was performed at Kolbe Mikroanalytisches Laboratorium using an ICP-optical emission spectrometer (Perkin Elmer) after sample dissolution according to standard in-house procedures.

**Electron microscopy studies.** HAADF imaging and EDX analysis (Figs 2 and 3, Extended Data Fig. 2) were performed on a Talos<sup>TM</sup> F200X transmission electron microscope (from FEI), equipped with a high-brightness field emission gun (X-FEG) and a Super-X G2 EDX detector. For these analyses, catalysts were embedded in Epofix resin, left to cure in air overnight at 60 °C, and cut to 70-nm sections using a Reichert-Jung Ultracut E ultramicrotome with Diatome Ultra 35° diamond knife. Sections were deposited on carbon-coated copper TEM grids. Images and elemental EDX maps were acquired using Velox<sup>TM</sup> analytical and imaging software in a scanning transmission mode, with a camera length of 77 mm. Elemental EDX maps of  $700 \times 700$  pixels (Fig. 2) were acquired with a 30-min acquisition time. Elemental EDX maps of  $512 \times 512$  pixels (Fig. 3b, d and Extended Data Fig. 2) were acquired with a 15-min acquisition time and processed in Velox<sup>TM</sup> using the box filter ( $3 \times 3$ ). Electron tomography was performed using Tecnai 20 (Extended Data Fig. 3a) and Talos<sup>TM</sup> F200X (Extended Data Fig. 3b) transmission electron microscopes. For electron tomography, catalysts were ground, sonicated in ethanol and drop-casted on Quantifoil R2/1 Cu TEM grids, with a thin carbon film and 5-nm Au particles as fiducial markers. Series of bright-field TEM images were taken with either a bottom-mounted charge-coupled-device camera from TVIPS (Extended Data Fig. 3a) or a Ceta 16M camera (Extended Data Fig. 3b), over an angle range of  $\pm 76^\circ$  with a tilt increment of  $2^\circ$ . Tilt series were aligned using the IMOD software package<sup>31</sup> and by tracking 5-nm Au fiducial markers from the TEM grid. Aligned series were binned by a factor of 2 and reconstructed in IMOD<sup>31</sup> using the weighted back projection (WBP) algorithm. Resulting reconstructions had a voxel size of  $(0.36 \text{ nm})^3$ ,  $(0.34 \text{ nm})^3$ ,  $(0.28 \text{ nm})^3$  or  $(0.26 \text{ nm})^3$ , depending on the magnification used.

**Catalytic tests.** Catalytic tests were performed in continuous-flow fixed-bed reactors. *n*-Decane (Sigma) hydroconversion was performed in a high-throughput reactor<sup>32,33</sup>. The catalyst weight was 50 mg; the total pressure was 0.45 MPa; the molar ratio of  $\text{H}_2$  to *n*-decane was 214; and the space time at the reactor entrance,  $W/F$  (contact time/rate of gas circulation), was  $1,400 \text{ kg s mol}^{-1}$ . *n*-Nonadecane (Sigma, 98%) was dissolved in heptane (Acros Organics) at 1 mol% concentration. The mixture was vaporized at 280 °C and mixed with hydrogen before feeding it to the reactor. The catalyst weight was 200 mg; the reaction was performed at a pressure of 0.65 MPa; the molar ratio of  $\text{H}_2$  to hydrocarbons was 14.6; and  $W/F$  was  $713 \text{ kg s mol}^{-1}$ . Under the investigated reaction conditions, heptane was confirmed to be inert. The same reaction conditions as for *n*-nonadecane conversion were used for pristane (Sigma, 98%) hydroconversion. The reaction products were analysed by on-line gas chromatography over an apolar capillary column (HP-1) and flame ionization detector. In both reactors, temperature was increased stepwise, and the reaction was equilibrated for 1 hour before product sampling.

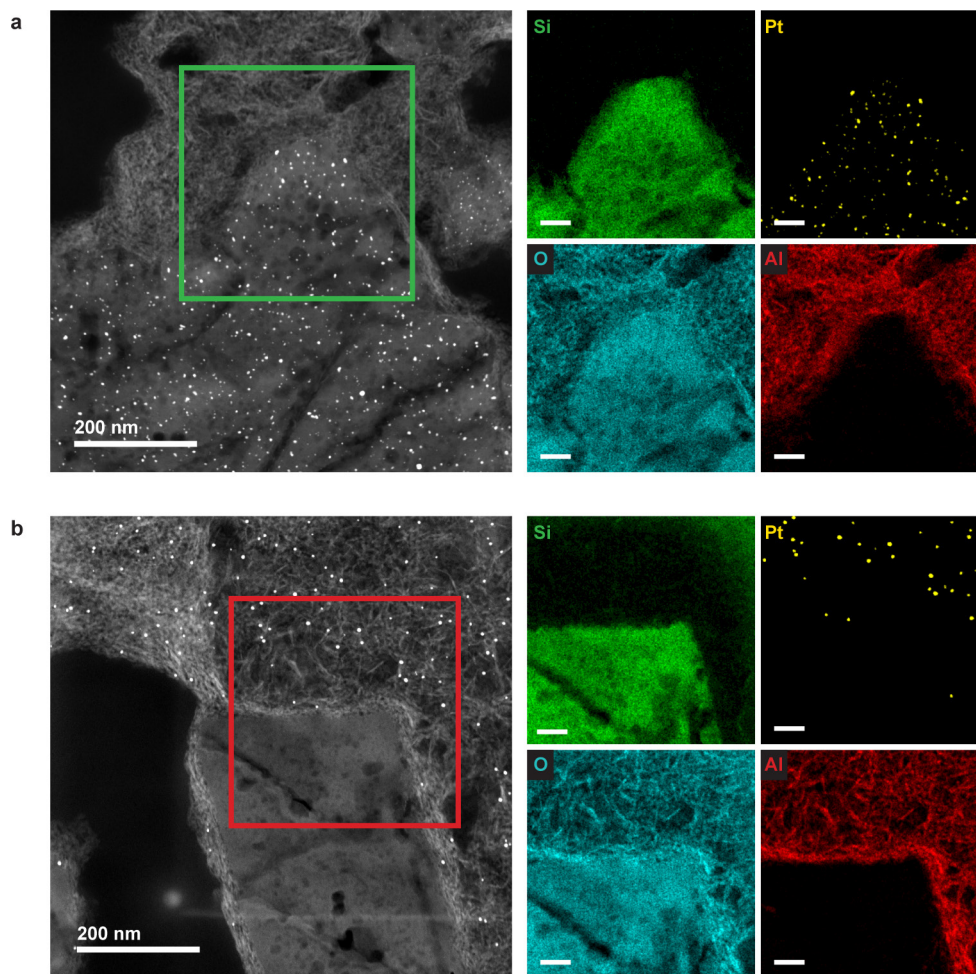
1. Kremer, J. R., Mastrorade, D. N. & McIntosh, J. R. Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol.* **116**, 71–76 (1996).
2. Huybrechts, W., Mijoin, J., Jacobs, P. A. & Martens, J. A. Development of a fixed-bed continuous-flow high-throughput reactor for long-chain *n*-alkane hydroconversion. *Appl. Catal. A* **243**, 1–13 (2003).
3. Burnens, G., Bouchy, C., Guillon, E. & Martens, J. Hydrocracking reaction pathways of 2,6,10,14-tetramethylpentadecane model molecule on bifunctional silica-alumina and ultrastable Y zeolite catalysts. *J. Catal.* **282**, 145–154 (2011).



#### Extended Data Figure 1 | Textural analysis of Y/A extrudates.

**a**, N<sub>2</sub> physisorption isotherm (adsorption and desorption), plotting the adsorbed N<sub>2</sub> volume (in cm<sup>3</sup> g<sup>-1</sup>), referred to a standard temperature and pressure (STP) of 0 °C and 1 atmosphere, against the relative pressure, P/P<sub>0</sub>. The isotherm shows the hysteresis loop that is indicative of the presence of mesopores (of diameter 2–50 nm). In addition, the quantity of

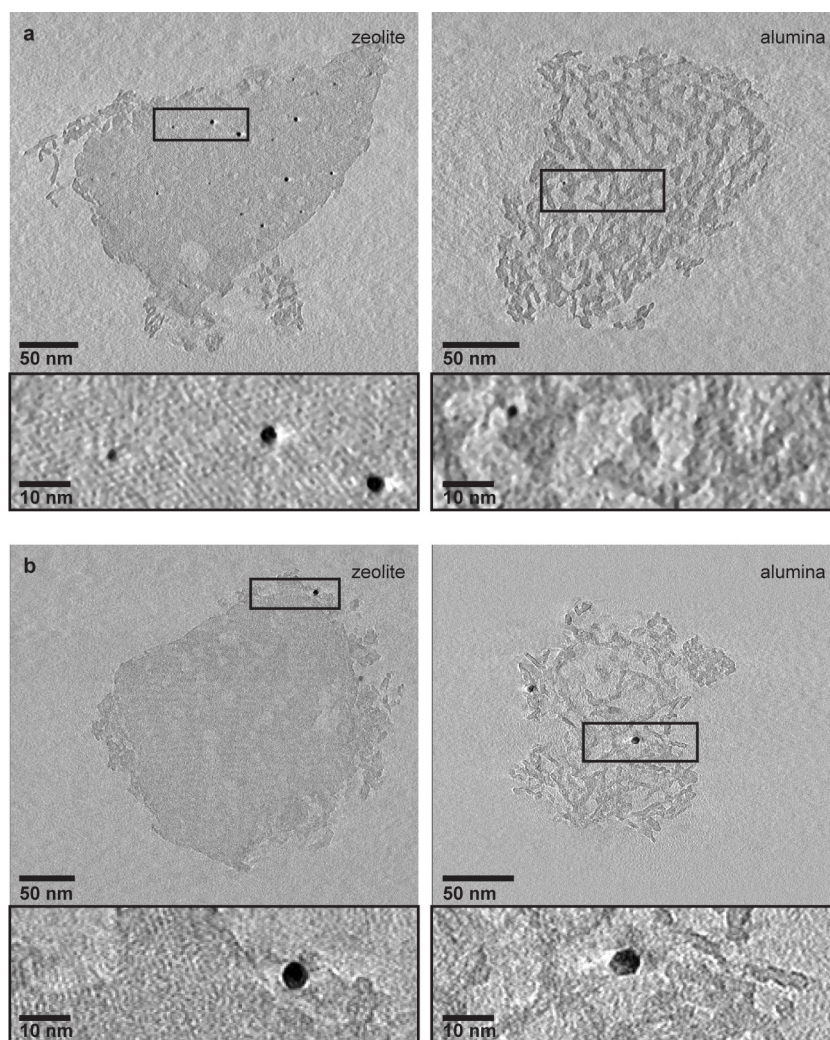
N<sub>2</sub> that is adsorbed at very low relative pressure indicates the presence of micropores (of diameter ~1 nm). **b**, The Barrett–Joyner–Halenda (BJH) pore-size distribution, derived from the adsorption branch of the isotherm in **a**, points to the broad size distribution of mesopores and macropores (macropores having a diameter greater than 50 nm). V is the pore volume; D is the pore diameter.



**Extended Data Figure 2 | Controlled deposition of platinum.** **a**, Left, an HAADF-STEM image of a 70-nm-thick section of the Pt-Y/A catalyst. Right, EDX elemental maps of the region outlined in green in the HAADF-STEM image. The EDX maps show the presence of Pt particles (yellow) in the zeolite region (green, with the dominant Si signal); the alumina region (red) is empty. 'O' denotes the oxygen signal. **b**, Left, an HAADF-STEM image of a 70-nm-thick section of the Pt-A/Y catalyst.

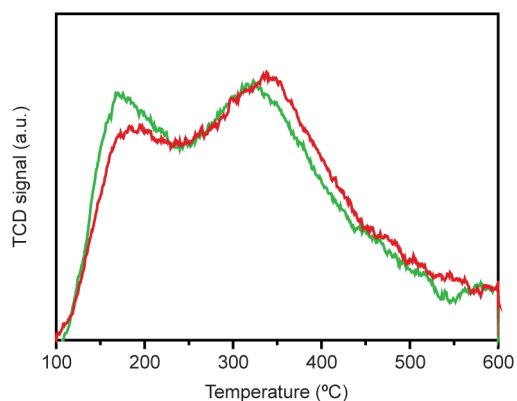
Right, EDX elemental maps of the region outlined in red at the left. Pt particles (yellow) are present in the alumina region (red), while the zeolite crystal (green) contains no Pt particles. The bright spot and line in the HAADF-STEM image of Pt-A/Y, below the mapped region of interest, originate from prolonged electron-beam exposure. Scale bars in the EDX maps represent 50 nm.





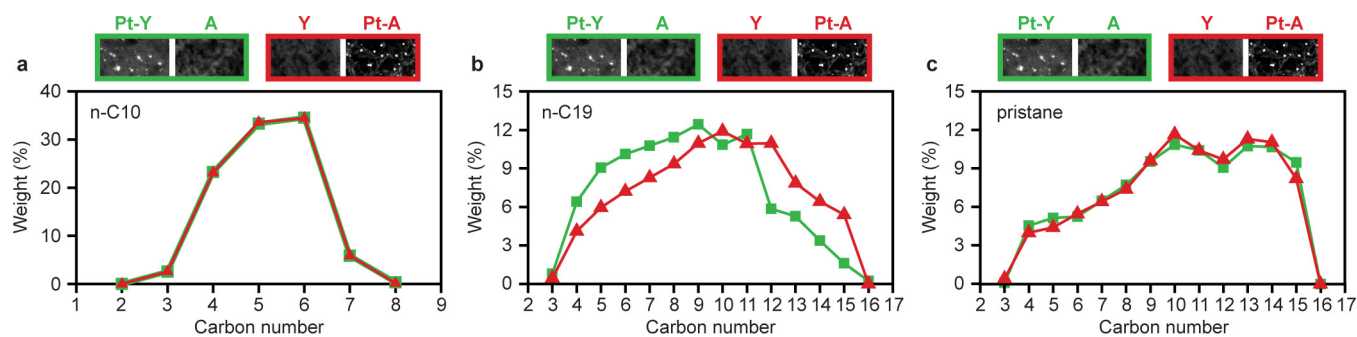
**Extended Data Figure 3 | Three-dimensional structural analysis using electron tomography. a**, Pt-Y/A catalyst. One-pixel slices, equal to thicknesses of 0.34 nm (left) and 0.26 nm (right), from the middle of the electron-tomography reconstructions of zeolite crystal (left) and alumina aggregate (right) from the Pt-Y/A catalyst, show the presence of  $\sim 2.5$ -nm Pt particles inside the zeolite crystal (left). Within the alumina aggregate (right), only a very few Pt particles were detected, of which one is shown in the zoomed-in region. **b**, Pt-A/Y catalyst. Ten-pixel slices, equal to thicknesses of 3.6 nm (left) and 2.8 nm (right), from the middle of the

electron-tomography reconstructions of zeolite crystal (left) and alumina aggregate (right) from the Pt-A/Y catalyst, show that Pt particles of  $\sim 3.5$ -nm diameter were located on the alumina platelets surrounding the zeolite crystal (left) and on the alumina platelets of the aggregate (right). No Pt particles were detected inside the zeolite crystal. For electron-tomography analysis, both catalysts were ground, dispersed in ethanol, and sonicated in order to break zeolite crystals and alumina aggregates apart and analyse them separately on the TEM grid.



**Extended Data Figure 4 | Comparing the acidity of the samples.**

Temperature-programmed desorption of ammonia by the Pt-Y/A (green) and Pt-A/Y (red) catalysts displays peaks that indicate the presence of weakly acidic (at  $\sim 160^\circ\text{C}$ ) and strongly acidic (at  $\sim 320^\circ\text{C}$ ) sites within both catalysts, with Pt-A/Y showing slightly lower peak intensity in the region of the weakly acidic site. The total amount of ammonia desorbed was measured to be  $14.3\text{ cm}^3\text{ g}^{-1}$  (Pt-Y/A) and  $14.7\text{ cm}^3\text{ g}^{-1}$  (Pt-A/Y) STP.



**Extended Data Figure 5 | Hydrocracking product distribution at 35% cracking conversion.** **a**, *n*-Decane feed; **b**, *n*-nonadecane feed; **c**, pristane feed. Results obtained with Pt-Y/A and Pt-A/Y catalysts are represented with green squares and red triangles, respectively. Experiments were

performed at a pressure of 0.45 MPa and H<sub>2</sub>/hydrocarbon molar ratio of 214 (*n*-decane); or a pressure of 0.65 MPa and a H<sub>2</sub>/hydrocarbon molar ratio of 14.6 (*n*-nonadecane and pristane).



# An observational radiative constraint on hydrologic cycle intensification

Anthony M. DeAngelis<sup>1</sup>, Xin Qu<sup>1</sup>, Mark D. Zelinka<sup>2</sup> & Alex Hall<sup>1</sup>

**Intensification of the hydrologic cycle is a key dimension of climate change, with substantial impacts on human and natural systems<sup>1,2</sup>. A basic measure of hydrologic cycle intensification is the increase in global-mean precipitation per unit surface warming, which varies by a factor of three in current-generation climate models (about 1–3 per cent per kelvin)<sup>3–5</sup>. Part of the uncertainty may originate from atmosphere–radiation interactions. As the climate warms, increases in shortwave absorption from atmospheric moistening will suppress the precipitation increase. This occurs through a reduction of the latent heating increase required to maintain a balanced atmospheric energy budget<sup>6,7</sup>. Using an ensemble of climate models, here we show that such models tend to underestimate the sensitivity of solar absorption to variations in atmospheric water vapour, leading to an underestimation in the shortwave absorption increase and an overestimation in the precipitation increase. This sensitivity also varies considerably among models due to differences in radiative transfer parameterizations, explaining a substantial portion of model spread in the precipitation response. Consequently, attaining accurate shortwave absorption responses through improvements to the radiative transfer schemes could reduce the spread in the predicted global precipitation increase per degree warming for the end of the twenty-first century by about 35 per cent, and reduce the estimated ensemble-mean increase in this quantity by almost 40 per cent.**

Although local precipitation changes are relevant for direct societal impacts, the change in globally averaged precipitation is a first-order indicator of climate change that must be well understood and accurately simulated. Evaluation of model-simulated global precipitation change with actual precipitation observations is difficult owing to uncertainties and insufficient spatial and temporal data coverage<sup>8</sup>. An alternative approach is to understand and evaluate the change in global precipitation through consideration of the physical processes that govern it.

Projected global-mean precipitation changes are dictated by the atmospheric energy budget's response to imposed radiative forcing and subsequent surface and atmospheric changes<sup>5–7,9,10</sup>. On annual and longer timescales, net atmospheric longwave cooling to the surface and outer space (LWC) is balanced by heating from shortwave absorption (SWA), sensible heating from the surface (SH), and latent heat release from precipitation ( $L_vP$ , latent heat of vaporization multiplied by precipitation rate) according to the following<sup>11</sup> (see also Fig. 1):

$$\text{LWC} = L_vP + \text{SWA} + \text{SH} \quad (1)$$

When carbon dioxide (CO<sub>2</sub>) and other greenhouse gases increase, and the planet warms, these energy sources and sinks readjust owing to a series of surface and atmospheric changes. Upon reaching equilibrium, the new balance is characterized by enhanced longwave cooling, increased solar absorption, decreased sensible heat flux from the surface to atmosphere, and increased precipitation<sup>6,11–13</sup> (Fig. 1b).

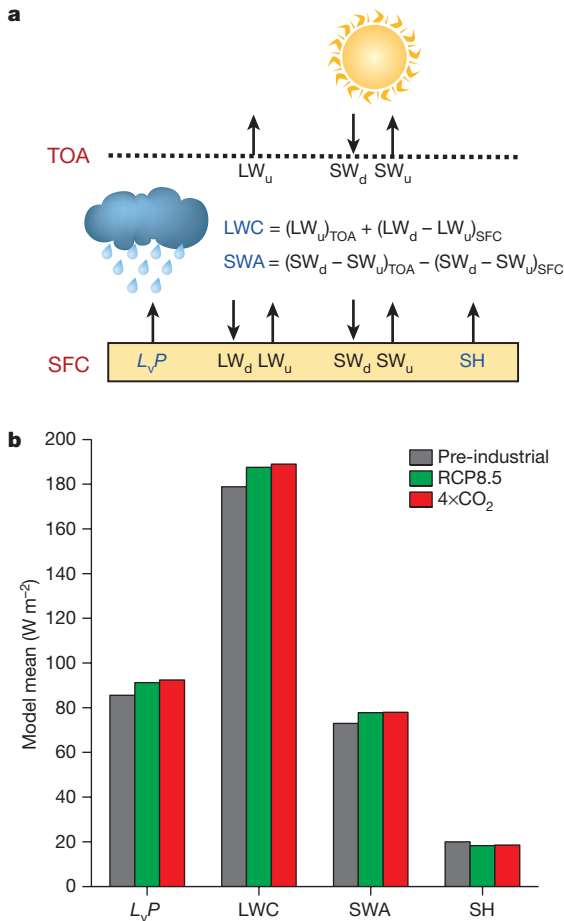
Climate model differences in the relative changes in the radiative and SH terms, per unit surface warming, produce different global-mean precipitation responses<sup>7,12,14</sup>. While it is clear that SWA increases mainly because of Clausius–Clayperon-driven increases in atmospheric water vapour, previous studies have demonstrated that the SWA response to warming has a notable spread across models<sup>4,7,12</sup>. Moreover, the physical basis for this spread and extent to which it can explain spread in the precipitation response remains a topic of debate<sup>4,7,12</sup>.

Global-mean precipitation responds to CO<sub>2</sub> forcing on two timescales<sup>9,15</sup>, as illustrated by idealized experiments where CO<sub>2</sub> suddenly increases. Precipitation initially decreases owing to a rapid increase in atmospheric static stability resulting from suppressed longwave cooling (the rapid adjustment)<sup>16,17</sup>, then slowly increases with subsequent global-mean surface warming (the temperature-mediated response)<sup>18,19</sup>. Previous studies investigating intermodel spread in the global-mean precipitation response to increased CO<sub>2</sub> have either not considered rapid adjustments and temperature-mediated responses separately (that is, they analysed total precipitation changes)<sup>4,12</sup>, or analysed a relatively small number of models, thereby undersampling the ensemble<sup>7</sup>. These limitations have hindered understanding of the sources of spread in hydrologic cycle intensification, particularly regarding the role of shortwave absorption<sup>4,7</sup>. Here we analyse rapid and temperature-mediated responses of global precipitation to CO<sub>2</sub> forcing in 25 models participating in the Coupled Model Intercomparison Project Phase 5 (CMIP5) to understand the spread.

Rapid adjustments and temperature-mediated responses are separated by regressing globally averaged annual anomalies in the atmospheric energy budget terms of equation (1) against surface air temperature ( $T$ ) anomalies in simulations of instantaneous CO<sub>2</sub> quadrupling (Methods). For each term, the regression slope represents the temperature-mediated response (hereafter  $dX/dT$ , where  $X$  is  $L_vP$ , SWA, LWC or SH) and the  $y$ -intercept where  $\Delta T = 0$  is the rapid adjustment<sup>20</sup> (Extended Data Fig. 1). The temperature-mediated precipitation response,  $L_v dP/dT$ , exhibits substantial intermodel spread, with values ranging from  $\sim 1.8$  to  $2.7 \text{ W m}^{-2} \text{ K}^{-1}$  ( $2.0\text{--}3.2\% \text{ K}^{-1}$ ) (Fig. 2a). This would contribute a  $4.5 \text{ W m}^{-2}$  spread in total  $L_vP$  change, the difference between a 10% and a 16% increase, if all models were to warm by  $\sim 5 \text{ K}$  (the multi-model mean warming averaged  $\sim 140\text{--}150$  years after CO<sub>2</sub> quadrupling). The spread in total  $L_vP$  change resulting from the rapid adjustment from the same forcing is smaller ( $\sim 3 \text{ W m}^{-2}$ ) (Extended Data Fig. 3). In this work, we focus on understanding spread in the temperature-mediated component.

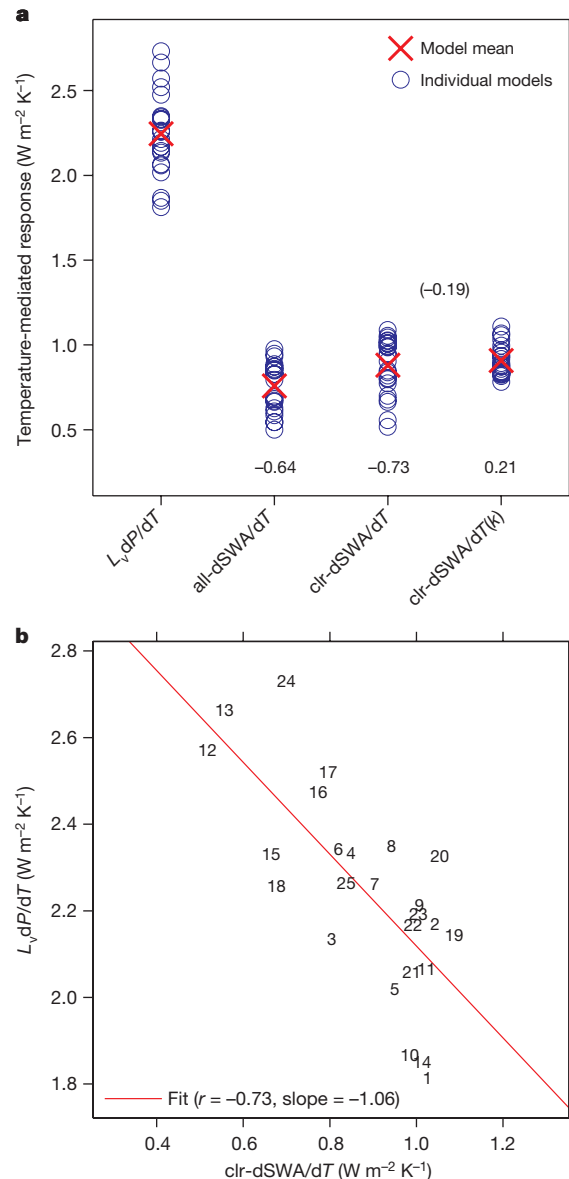
While longwave cooling is the key driver of temperature-mediated precipitation change in the multi-model mean, the spread in  $d\text{LWC}/dT$  and  $d\text{SWA}/dT$  is large, and each accounts for substantial intermodel spread in  $L_v dP/dT$  (the correlation coefficient,  $r$ , between each and  $L_v dP/dT$  exceeds magnitude 0.60) (Extended Data Fig. 2). The anti-correlation between  $d\text{SWA}/dT$  and  $L_v dP/dT$  is somewhat larger in magnitude and occurs in both all-sky ( $r = -0.64$ ) and clear-sky ( $r = -0.73$ )

<sup>1</sup>Department of Atmospheric and Oceanic Sciences, University of California Los Angeles, Los Angeles, California 90095, USA. <sup>2</sup>Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, California 94550, USA.



**Figure 1 | The atmospheric energy budget.** **a**, Schematic of the fluxes considered in the budget. All fluxes (arrows) shown in the schematic are defined positive; some represent downward (denoted with subscript d) and others upward (denoted with subscript u) fluxes at the top-of-atmosphere (TOA) or surface (SFC). All flux arrows are drawn with the same length for simplicity. The four main energy budget terms considered in this study (equation (1) in the main text) are identified with blue text here. Two of them are radiative and are derived from individual longwave (LW) or shortwave (SW) flux components, following the equations shown here: net atmospheric longwave cooling (LWC, the sum of upward LW flux at TOA and net downward LW flux at the SFC); and shortwave absorption (SWA, net downward SW flux at TOA minus net downward SW flux at the SFC). The other two are the latent heat release from precipitation ( $L_vP$ , latent heat of vaporization multiplied by precipitation rate) and sensible heat flux from the surface to atmosphere (SH). **b**, Grey, the global-mean values of the four terms simulated by the CMIP5 ensemble for the pre-industrial control simulation (average over 150 yr); green, Representative Concentration Pathway scenario 8.5 (RCP8.5, average over 2091–2100); and red, abrupt quadrupled  $CO_2$  scenario (average for years 141–150 after  $CO_2$  quadrupling in all models except CNRM-CM5-2 and IPSL-CM5A-MR, where it is for years 131–140). The model mean is shown in all cases, and the radiative terms are computed using all-sky fluxes. All terms are positive.

conditions (Fig. 2a). This suggests clouds play a negligible role in the relationship<sup>7,10</sup> (Extended Data Fig. 2b). An anti-correlation arises because SWA and  $L_vP$  compete to balance enhanced longwave cooling in a warmer climate (Fig. 1). Thus models with a larger SWA increase tend to have a smaller  $L_vP$  increase, per unit surface warming. The importance of solar absorption for model spread in the global-mean precipitation response to increased  $CO_2$  was also demonstrated in ref. 7; however, we find that the importance of solar absorption stems mainly from the temperature-mediated component of the change, which is inconsistent with the findings in ref. 7 (Methods). Based on our results,  $dSWA/dT$  is a key source of model spread in hydrologic



**Figure 2 | Relationship between temperature-mediated  $L_vP$  and SWA responses.** **a**, The temperature-mediated responses of  $L_vP$  (that is,  $L_vP/dT$ ) and SWA (that is,  $dSWA/dT$ ) for all-sky (indicated with all-) and clear-sky (indicated with clr-) conditions are shown. The values are computed using model-produced fluxes (as in Extended Data Fig. 1) except in the case of  $clr-dSWA/dT(k)$ , in which they are computed with radiative kernels (Methods). Individual CMIP5 models are shown as blue circles. The multi-model mean is identified with a red cross. The numbers above the abscissa are the cross-model correlations between  $L_vP/dT$  and each  $dSWA/dT$ . The number in parentheses is the correlation between model-produced and kernel-derived clear-sky  $dSWA/dT$ . **b**, Scatterplot of the model-produced  $L_vP/dT$  versus clear-sky  $dSWA/dT$  for the 25 models, with corresponding least-squares linear fit shown (red line). Note that the abscissa and ordinate axes have the same scale. Model numbers are defined in Extended Data Table 1.

cycle intensification, analogous to the importance of cloud feedbacks for model uncertainty in climate sensitivity.

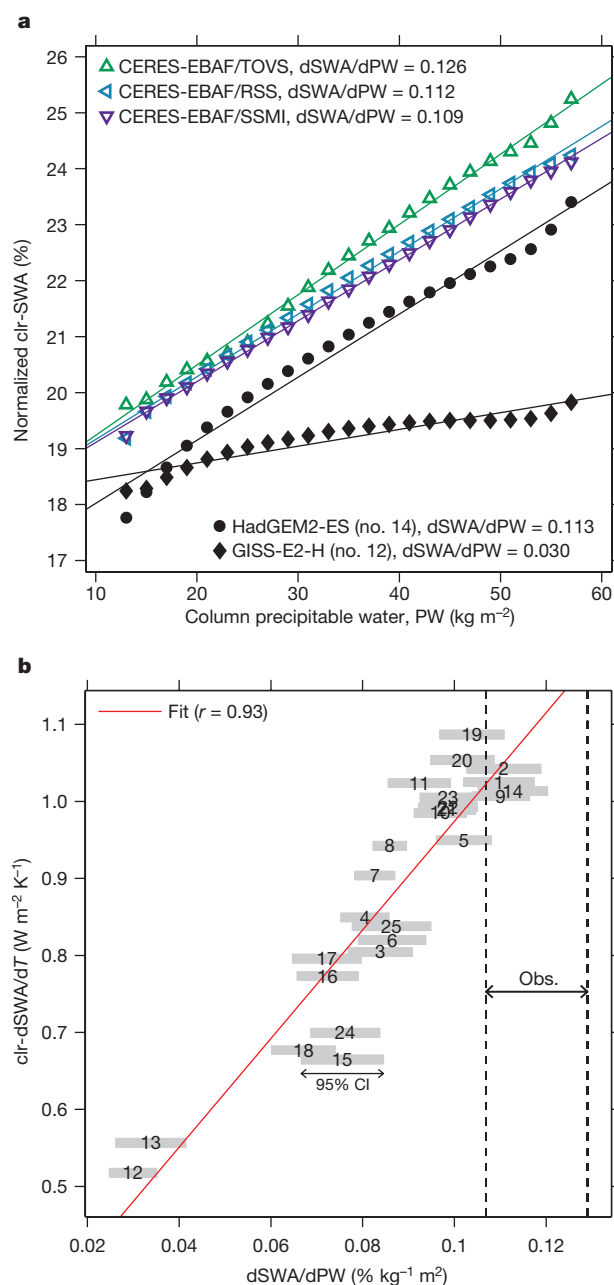
What generates the spread in  $dSWA/dT$ ? Under greenhouse-gas induced warming, SWA increases mainly because of enhanced solar absorption by water vapour in a warmer and moister atmosphere<sup>7,12</sup>. Thus, model spread in  $dSWA/dT$  arises from differences in water vapour absorption. This is supported by the large anti-correlation between  $L_vP/dT$  and clear-sky  $dSWA/dT$  (Fig. 2b). Previous studies have proposed two potential sources of model disagreement in

the simulated increase of solar absorption: (1) different increases in global-mean column water vapour and/or different vertical/horizontal patterns of vapour change for the same surface warming<sup>7</sup>; and (2) differences in the sensitivity of solar absorption to a unit change in atmospheric water vapour content, related to radiative transfer parameterizations<sup>12</sup>. In this Letter, we seek to determine which source is more important and to better understand its physical basis. We focus on clear-sky SWA, and hereafter, SWA will refer to its clear-sky component.

Does model spread in  $d\text{SWA}/dT$  come from the first source, that is, differences in the water vapour response for the same surface warming? Cross-model correlations between globally averaged temperature-mediated water vapour responses (both column-integrated and at individual levels) and  $d\text{SWA}/dT$  are weak and statistically insignificant, suggesting that it does not (Methods, Extended Data Fig. 4). To substantiate this further and to examine the role of horizontal variability in vapour changes, we estimate  $d\text{SWA}/dT$  using radiative kernels (linear approximations of radiative flux sensitivity to perturbations in atmospheric state; Methods). As the kernels are developed with a single radiative transfer code<sup>14</sup>, model spread in kernel-derived  $d\text{SWA}/dT$  mainly reflects intermodel differences in the water vapour response to surface warming<sup>14</sup>. If the vapour response were causing most of the spread in  $d\text{SWA}/dT$ , the kernel-derived  $d\text{SWA}/dT$  and actual  $d\text{SWA}/dT$  computed from model-produced fluxes should be highly correlated, with a similar spread. However, kernel-derived  $d\text{SWA}/dT$  exhibits approximately half the spread and is uncorrelated with model-produced  $d\text{SWA}/dT$  (Fig. 2a). Thus model variations in  $d\text{SWA}/dT$  are not accounted for by the total water vapour increase or by subtle differences in the vertical/horizontal structure of vapour increase among models. This is consistent with the findings of ref. 12. Instead, the main source of intermodel variability in  $d\text{SWA}/dT$  would appear to be the sensitivity of solar absorption to a unit change in atmospheric water vapour.

The sensitivity of solar absorption to a change in water vapour concentration is estimated in each model by binning SWA based on column precipitable water (PW) taken from each grid cell and month in the control climate and computing the regression slope of bin-averaged SWA against PW (Methods). This quantity is hereafter denoted  $d\text{SWA}/d\text{PW}$ . Figure 3a shows SWA versus PW curves for the two models having the largest and smallest sensitivity.  $d\text{SWA}/d\text{PW}$  is in the range  $0.03\text{--}0.11\% \text{ kg}^{-1} \text{ m}^2$  across models (Fig. 3), with much of the difference arising from absorption under moist conditions (Extended Data Fig. 5).  $d\text{SWA}/d\text{PW}$  is almost perfectly correlated with  $d\text{SWA}/dT$  across models ( $r = 0.93$ , Fig. 3b). This confirms that the sensitivity of solar absorption to a change in atmospheric moisture is the principal source of spread in the temperature-mediated SWA response to  $\text{CO}_2$  forcing.

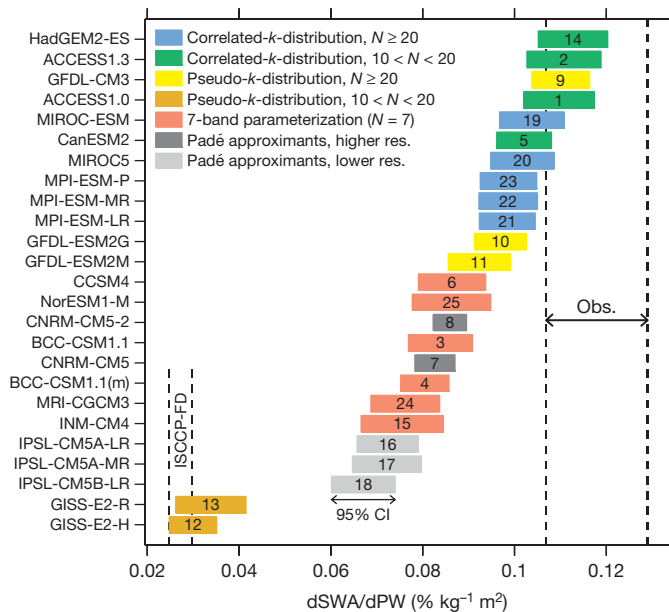
We estimate the real atmosphere's  $d\text{SWA}/d\text{PW}$  by combining radiative fluxes from the CERES-EBAF (Clouds and the Earth's Radiant Energy System Energy Balance and Filled) products<sup>21,22</sup> with PW measurements from three sources (Methods).  $d\text{SWA}/d\text{PW}$  obtained from CERES-EBAF and each PW data set ( $\sim 0.11\text{--}0.13\% \text{ kg}^{-1} \text{ m}^2$ ) is on the upper end of the CMIP5 range (Fig. 3), implying that  $d\text{SWA}/d\text{PW}$  is too weak in most models. This is consistent with recent studies showing that many climate models underestimate solar absorption in moist atmospheres<sup>23–25</sup>. Thus the temperature-mediated SWA response is underestimated by many models. Models with  $d\text{SWA}/d\text{PW}$  within statistical uncertainty of observations have  $d\text{SWA}/dT$  in the range  $0.95\text{--}1.09 \text{ W m}^{-2} \text{ K}^{-1}$ , that is, the upper 25% of the full spread of  $0.52\text{--}1.09 \text{ W m}^{-2} \text{ K}^{-1}$  (Fig. 3b). Given the strong anti-correlation between  $d\text{SWA}/dT$  and  $L_v dP/dT$  (Fig. 2b), this further implies that many models overestimate the temperature-mediated increase in global-mean precipitation, assuming no compensating errors in other energy budget terms. This assumption may not be true, however, as models may underestimate projected increases in longwave cooling because of a hypothesized missing iris effect<sup>26</sup>. If this is the case, it would contribute an underestimation in global precipitation increase that is independent of temperature-mediated shortwave absorption (Methods).



**Figure 3 | The sensitivity of solar absorption to varying atmospheric water vapour.** **a**, Clear-sky (clr)-SWA normalized by incoming solar flux versus column precipitable water (PW) in the pre-industrial climate for selected models and for CERES-EBAF estimates with three PW data sets (see Methods). The slope of the curve represents the sensitivity of SWA to varying PW ( $d\text{SWA}/d\text{PW}$ ) and its value is given in the keys for each model/data set (in  $\% \text{ kg}^{-1} \text{ m}^2$ ). **b**, Scatterplot of the (clear-sky) temperature-mediated SWA response to  $\text{CO}_2$  forcing,  $d\text{SWA}/dT$  (Fig. 2b), versus  $d\text{SWA}/d\text{PW}$  (**a**) for the 25 models, with corresponding linear fit shown. In **b**, the width of the horizontal shading for each model represents the 95% confidence interval (CI) of the regression slope to the SWA versus PW curve (Methods). Statistical uncertainty of the CERES-EBAF  $d\text{SWA}/d\text{PW}$  estimate combining all PW data sets is shown with vertical dashed lines (Obs.). It is computed by (1) calculating the 95% CI of the slope of the CERES-EBAF curve with each individual PW data set, then (2) calculating the absolute minimum to maximum of the three CIs (that is, the minimum of the lower bounds of all CIs to the maximum of the upper bounds of all CIs).

What is the physical basis for the model spread in  $d\text{SWA}/d\text{PW}$ ? Simulated solar absorption for a given atmospheric water vapour concentration is determined by radiative transfer algorithms that approximate complex spectral absorption by water vapour molecules, among





**Figure 4 | Shortwave parameterization schemes.** The relationship between dSWA/dPW (Fig. 3) and characteristics of the parameterization scheme for solar absorption by water vapour in a cloud-free atmosphere, with colours for each model referring to different types of parameterizations as described in the key ( $N$  refers to the number of exponential terms representing water vapour absorption). References and further discussion are given in Extended Data Table 1 and Methods, respectively. Model numbers are identified. The width of horizontal shading for models and the vertical dashed lines for observations (Obs.) represent statistical uncertainties of dSWA/dPW, as described in Fig. 3. The 95% CI for an estimate of dSWA/dPW based on ISCCP-FD (Methods) is also shown with vertical dashed lines.

other constituents. We examine the influence of these algorithms on simulated dSWA/dPW by categorizing the shortwave parameterization scheme in each model based on its treatment of clear-sky solar absorption by water vapour (Fig. 4). The schemes vary considerably among models and their characteristics have a strong correspondence with dSWA/dPW. In general, models that implement more modern approaches and/or use a larger number of mathematical terms ( $N$  in Fig. 4) to approximate the complex dependency of shortwave transmission on wavelength tend to have larger and more realistic dSWA/dPW (see also Methods). It is important to note that a commonly used observational product, the International Satellite Cloud Climatology Project flux data set (ISCCP-FD)<sup>27</sup>, exhibits a large bias in dSWA/dPW (Fig. 4). This product generates fluxes with a radiative transfer algorithm nearly identical to that used in the GISS-E2-H and GISS-E2-R models (Methods), and as a consequence yields the same dSWA/dPW as those models. Thus, caution should be exercised when treating ISCCP-FD radiative fluxes as observations.

Intermodel variability in the accuracy of shortwave parameterization schemes probably results from model developers' ongoing challenge of balancing the need for accurate radiative transfer calculations against considerations of computational efficiency and realistic simulation of other climate system components<sup>24</sup>. As computational capabilities have grown, improvement in longwave schemes and other model components (for example, cloud processes) seem to have taken precedence over parameterization of shortwave gaseous absorption, with many modelling institutions continuing to implement outdated schemes for the latter in CMIP5<sup>24,25,28</sup> (Extended Data Fig. 6). This is understandable considering the importance of cloud feedbacks and longwave fluxes and for climate sensitivity, but here we show that atmospheric solar absorption is equally important for hydrologic cycle intensification. Based on a simple calculation (Methods), we estimate that if the temperature-mediated SWA response to CO<sub>2</sub>

forcing were perfectly constrained in the current generation of models, the spread in total precipitation increase per unit warming predicted at the end of the twenty-first century under the Representative Concentration Pathway scenario 8.5 (RCP8.5)<sup>29</sup> could be reduced by ~35%, and the ensemble-mean precipitation increase per unit warming would decrease by nearly 40% (Extended Data Fig. 7). Even for the total precipitation change (that is, not normalized by surface warming), a spread reduction of ~25% and reduction in ensemble-mean increase of ~25% could be obtained. Clearly, shortwave radiative transfer parameterizations need improvement in many climate models, and fortunately, this is already under way at several modelling institutions<sup>25</sup>.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 26 May; accepted 25 September 2015.**

- Trenberth, K. E. Changes in precipitation with climate change. *Clim. Res.* **47**, 123–138 (2011).
- Pendergrass, A. G. & Hartmann, D. L. Changes in the distribution of rain frequency and intensity in response to global warming. *J. Clim.* **27**, 8372–8383 (2014).
- Kharin, V. V., Zwiers, F. W., Zhang, X. & Wehner, M. Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Clim. Change* **119**, 345–357 (2013).
- Pendergrass, A. G. & Hartmann, D. L. Global-mean precipitation and black carbon in AR4 simulations. *Geophys. Res. Lett.* **39**, L01703 (2012).
- Andrews, T. & Forster, P. M. The transient response of global-mean precipitation to increasing carbon dioxide levels. *Environ. Res. Lett.* **5**, 025212 (2010).
- O'Gorman, P. A., Allan, R. P., Byrne, M. P. & Previdi, M. Energetic constraints on precipitation under climate change. *Surv. Geophys.* **33**, 585–608 (2012).
- Takahashi, K. The global hydrological cycle and atmospheric shortwave absorption in climate models under CO<sub>2</sub> forcing. *J. Clim.* **22**, 5667–5675 (2009).
- Arkin, P. A., Smith, T. M., Sapiano, M. R. P. & Janowiak, J. The observed sensitivity of the global hydrological cycle to changes in surface temperature. *Environ. Res. Lett.* **5**, 035201 (2010).
- Allen, M. R. & Ingram, W. J. Constraints on future changes in climate and the hydrologic cycle. *Nature* **419**, 224–232 (2002).
- Lambert, F. H. & Webb, M. J. Dependency of global mean precipitation on surface temperature. *Geophys. Res. Lett.* **35**, L16706 (2008).
- Manabe, S. & Wetherald, R. T. The effects of doubling the CO<sub>2</sub> concentration on the climate of a general circulation model. *J. Atmos. Sci.* **32**, 3–15 (1975).
- Pendergrass, A. G. & Hartmann, D. L. The atmospheric energy constraint on global-mean precipitation change. *J. Clim.* **27**, 757–768 (2014).
- Mitchell, F. B., Wilson, C. A. & Cunningham, W. M. On CO<sub>2</sub> climate sensitivity and model dependence of results. *Q. J. R. Meteorol. Soc.* **113**, 293–322 (1987).
- Previdi, M. Radiative feedbacks on global precipitation. *Environ. Res. Lett.* **5**, 025211 (2010).
- Yang, F., Kumar, A., Schlesinger, M. E. & Wang, W. Intensity of hydrological cycles in warmer climates. *J. Clim.* **16**, 2419–2423 (2003).
- Dong, B., Gregory, J. M. & Sutton, R. T. Understanding land–sea warming contrast in response to increasing greenhouse gases. Part I: Transient adjustment. *J. Clim.* **22**, 3079–3097 (2009).
- Bony, S. *et al.* Robust direct effect of carbon dioxide on tropical circulation and regional precipitation. *Nature Geosci.* **6**, 447–451 (2013).
- Bala, G., Caldeira, K. & Nemani, R. Fast versus slow response in climate change: implications for the global hydrological cycle. *Clim. Dyn.* **35**, 423–434 (2010).
- Andrews, T., Forster, P. M., Boucher, O., Bellouin, N. & Jones, A. Precipitation, radiative forcing and global temperature change. *Geophys. Res. Lett.* **37**, L14701 (2010).
- Gregory, J. M. *et al.* A new method for diagnosing radiative forcing and climate sensitivity. *Geophys. Res. Lett.* **31**, L03205 (2004).
- Wielicki, B. A. *et al.* Clouds and the Earth's Radiant Energy System (CERES): An Earth observing system experiment. *Bull. Am. Meteorol. Soc.* **77**, 853–868 (1996).
- Kato, S. *et al.* Surface irradiances consistent with CERES-derived top-of-atmosphere shortwave and longwave irradiances. *J. Clim.* **26**, 2719–2740 (2013).
- Collins, W. D. *et al.* Radiative forcing by well-mixed greenhouse gases: Estimates from climate models in the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4). *J. Geophys. Res.* **111**, D14317 (2006).
- Oreopoulos, L. *et al.* The continual intercomparison of radiation codes: results from phase I. *J. Geophys. Res.* **117**, D06118 (2012).
- Pincus, R. *et al.* Radiative flux and forcing parameterization error in aerosol-free clear skies. *Geophys. Res. Lett.* **42**, 5485–5492 (2015).
- Mauritsen, T. & Stevens, B. Missing iris effect as a possible cause of muted hydrological change and high climate sensitivity in models. *Nature Geosci.* **8**, 346–351 (2015).

27. Zhang, Y., Rossow, W. B., Lacis, A. A., Oinas, V. & Mishchenko, M. I. Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: Refinements of the radiative transfer model and the input data. *J. Geophys. Res.* **109**, D19105 (2004).
28. Voldoire, A. *et al.* The CNRM-CM5.1 global climate model: description and basic evaluation. *Clim. Dyn.* **40**, 2091–2121 (2013).
29. Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview of CMIP5 and the experimental design. *Bull. Am. Meteorol. Soc.* **93**, 485–498 (2012).

**Acknowledgements** All authors were supported by the Regional and Global Climate Modeling Program of the Office of Science of the US Department of Energy. The work of M.D.Z. was performed under the auspices of the US Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modelling groups (listed in Extended Data Table 1) for producing and making available their model output. For CMIP, the US Department of Energy's Program for Climate Model Diagnosis and Intercomparison provided coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. We acknowledge use of the CERES-EBAF flux data

obtained from the National Aeronautics and Space Administration (NASA) Langley Research Center ([http://ceres.larc.nasa.gov/order\\_data.php](http://ceres.larc.nasa.gov/order_data.php)), the ISCCP-FD data obtained from the NASA Goddard Institute for Space Studies (<http://isccp.giss.nasa.gov/projects/flux.html>), the SSM/I data obtained from the National Oceanic and Atmospheric Administration (NOAA) National Climatic Data Center (<http://www.ncdc.noaa.gov/oa/rsad/ssmi/gridded/index.php>), and the RSS data obtained from Remote Sensing Systems (<http://www.remss.com/measurements/atmospheric-water-vapor/tpw-1-deg-product>). We thank M. Previdi for providing the radiative kernels. We also thank S. A. Klein, K. E. Taylor, P. M. Caldwell, A. A. Lacis, R. Pincus and A. J. Broccoli for discussion on the topic.

**Author Contributions** A.M.D., X.Q. and A.H. designed the methodology. A.M.D. performed the analysis and wrote the paper. M.D.Z. provided the kernel-derived temperature-mediated shortwave absorption response estimates. All authors discussed the results and edited the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.M.D. ([adeangelis@ucla.edu](mailto:adeangelis@ucla.edu)).

## METHODS

**CMIP5 models.** The CMIP5 models<sup>29</sup> analysed here are listed in Extended Data Table 1, with corresponding references. We use one ensemble member (r1i1p1) from 25 models that had available monthly output of atmospheric temperature and humidity, precipitation, and energy fluxes for the pre-industrial control (piControl) and abrupt quadrupled CO<sub>2</sub> (abrupt4 × CO<sub>2</sub>) experiments at the time of analysis. **Temperature-mediated responses and rapid adjustments.** The Gregory method<sup>10,20</sup> is employed using each model's piControl and abrupt4 × CO<sub>2</sub> run. For each year of the abrupt4 × CO<sub>2</sub> run, the global-mean abrupt4 × CO<sub>2</sub> anomaly of a physical quantity (for example,  $L_pP$ ) relative to the piControl simulation is paired against the corresponding anomaly of 2-m air temperature, generating a scatterplot (Extended Data Fig. 1). To compute the annual anomalies, the piControl 21-year mean, centred on the corresponding year of the abrupt4 × CO<sub>2</sub> simulation, is subtracted from each abrupt4 × CO<sub>2</sub> 1-year mean. Subtracting this running mean removes possible influences of climate model drift on the anomalies. The scatterplots are generated using 150 years of the abrupt4 × CO<sub>2</sub> simulation when available (140 years are available from the CNRM-CM-2 and IPSL-CM5A-MR models). A least-squares linear regression is then applied to each scatterplot, with the slope and  $y$ -intercept of the fit representing the temperature-mediated and rapid responses to CO<sub>2</sub> forcing, respectively.

The Gregory methodology is displayed visually for the GFDL-CM3 model in Extended Data Fig. 1. It shows the yearly evolution of globally averaged LWC, SWA, SH, and  $L_pP$  changes after a quadrupling of atmospheric CO<sub>2</sub>. The physical interpretation of these changes has been thoroughly discussed in numerous studies<sup>6,12,16,18,30–32</sup>. The high degree of linearity of the scatterplots demonstrates the reliability of this approach for separating temperature-mediated responses and rapid adjustments.

As shown in Fig. 2, model spread in the temperature-mediated SWA response,  $dSWA/dT$ , particularly for clear-sky, substantially contributes to the spread in  $L_pP/dT$ . The total change in  $L_pP$  per unit warming ( $L_p\Delta P/\Delta T$ , which includes the temperature-mediated response and rapid adjustment), however, is not correlated with the corresponding total change in SWA across models (Extended Data Fig. 3a). This may be the consequence of large scatter in the rapid adjustment of  $L_pP$ , which is not strongly related to that of SWA (Extended Data Fig. 3b). Reference 7 found the opposite result using eight CMIP3 models: that the temperature-mediated responses of latent heating and solar absorption are not correlated, but that total changes at a warming of 2 K are anti-correlated ( $r = -0.66$ ). The differences between the results in ref. 7 and ours probably originates from different analysis methods and sampling. For example, in ref. 7 only 8 models from CMIP3 were used, whereas we use 25 from CMIP5. Additionally, ref. 7 combined different types of CO<sub>2</sub> forcing scenarios (for example, 1% increase to CO<sub>2</sub> doubling or quadrupling and instantaneous CO<sub>2</sub> doubling) and based their temperature-mediated estimates on scatterplots of relatively few (5–10) data points from each simulation. It is shown in their analysis that quite different slopes may be obtained from different forcing simulations even with the same model. We argue that our results are more robust. In addition to using many more models, we base our temperature-mediated estimates on a large number of data points (140–150) from a single forcing scenario (instantaneous CO<sub>2</sub> quadrupling). Furthermore, our findings are consistent with results reported in ref. 4, where it was also found that total changes in precipitation (per unit warming) are not correlated with those in shortwave absorption under doubled CO<sub>2</sub> using a larger sample of CMIP3 models (14) than used in ref. 7.

**Radiative kernels and water vapour responses.** Radiative kernels were developed using an offline version of the MPI-ECHAM5 radiation code and represent the sensitivity of top-of-atmosphere (TOA) and surface shortwave radiative fluxes to small perturbations in atmospheric specific humidity and surface albedo<sup>14</sup>. Global-mean temperature-mediated SWA responses due to water vapour changes are computed as follows: (1) temperature-mediated responses in the logarithm of specific humidity at all months, locations and pressure levels are multiplied by the shortwave specific humidity atmospheric (TOA minus surface) kernel for clear-sky. Global-annual-mean kernel values are shown in Extended Data Fig. 4a. Specific humidity responses are computed as the difference in abrupt4 × CO<sub>2</sub>—piControl anomalies averaged over years 121–150 and 1–30 of the abrupt4 × CO<sub>2</sub> simulation, normalized by the corresponding difference in 2-m air temperature. (2) The product is integrated over the depth of the troposphere (defined as all levels between the surface and a tropopause height that varies linearly with latitude from 100 hPa on the Equator to 300 hPa at the poles), and then averaged over all months and locations. Temperature-mediated SWA responses due to surface albedo are computed in similar fashion but with surface albedo kernels and temperature-mediated responses. The final kernel-derived temperature-mediated SWA response for each model ( $dSWA/dT(k)$  in Fig. 2a) is the sum of the water vapour and surface albedo components of the response, with the vapour component dominating<sup>14</sup>.

It should be noted that the computation of atmospheric radiative feedbacks using kernels is associated with uncertainty<sup>12,14</sup>. In this study, our goal is not to

obtain perfectly accurate quantitative estimates of temperature-mediated responses with the kernels. Rather, we apply the kernels to assess whether model spread in  $dSWA/dT$  originates from differences in the response of atmospheric water vapour to surface warming. We argue that the kernel calculations are sufficiently accurate for this purpose. Extended Data Fig. 4 shows that kernel-derived  $dSWA/dT$  is significantly correlated across models with globally averaged responses of water vapour, as expected. By contrast, actual model-produced  $dSWA/dT$  is not correlated with the water vapour responses, supporting our finding that spread in  $dSWA/dT$  is not explained by differences in the water vapour response. This may be partly because water vapour responses (in a fractional sense) and their intermodel variability are largest in the upper troposphere, where water vapour changes have little effect on total-column SWA (Extended Data Fig. 4a). Our conclusions are also supported by ref. 12, where it is independently shown that CMIP5 spread in enhanced solar absorption under CO<sub>2</sub> forcing cannot be reproduced by imposing simulated changes in atmospheric temperature and moisture on the single-column Fu-Liou radiative transfer model. Nonetheless, to assess uncertainty in kernel-estimated temperature-mediated responses, calculations with additional kernels produced from different radiation codes would be necessary.

**The sensitivity of SWA to a unit water vapour change.** SWA sensitivity to water vapour variability is computed on the basis of spatial and temporal variations of PW in the control climate. All grid cells and months from 150 years of a model's piControl simulation over the tropical oceans are aggregated into one sample to compute the sensitivity. The years 2001–2009 and 1984–2009 are used for estimates based on the CERES-EBAF (Clouds and the Earth's Radiant Energy System Energy Balance and Filled) and ISCCP-FD (International Satellite Cloud Climatology Project flux data set) products, respectively. Tropical oceans are defined as grid cells with centres between 30° S and 30° N and with land fraction less than 0.50. All model output is regridded to  $2.5^\circ \times 2.5^\circ$  latitude–longitude before performing calculations. To compute the sensitivity, the clear-sky SWA at each grid cell and for each month is normalized by incoming solar radiation, then binned according to PW with equal bin size of  $2 \text{ kg m}^{-2}$ . SWA is averaged within each PW bin and plotted against the bin centre value (Fig. 3a). Only bins with at least 20 data values in every model and observational source are considered, resulting in a common PW range of 12–58  $\text{kg m}^{-2}$ . The linear regression slope of the SWA versus PW scatterplot represents the SWA sensitivity to varying PW (that is,  $dSWA/dPW$ ). Statistical uncertainty in  $dSWA/dPW$  is computed as the 95% confidence interval (CI) of the regression slope, derived from the estimated standard error of the slope parameter<sup>33</sup>.

We consider only tropical oceans when computing  $dSWA/dPW$  owing to the relatively small variability of surface albedo and solar zenith angle within this region, better isolating the effect of PW on SWA. To examine whether the remaining small variations of surface albedo or solar zenith angle affect  $dSWA/dPW$ , we re-compute  $dSWA/dPW$  by also conditioning on surface albedo and/or zenith angle. Specifically,  $dSWA/dPW$  is recomputed from locations and months where albedo and/or zenith angle vary by no more than 0.01 or  $1^\circ$ , respectively. While the recomputed  $dSWA/dPW$  values vary slightly from the original values computed with the entire domain, they are highly correlated with the original values across models (not shown). This suggests that the influence of albedo and zenith angle on  $dSWA/dPW$  is very small.

Our computation of  $dSWA/dPW$  does not isolate the effect of aerosols on SWA. Models differ in the types of aerosols represented, the concentrations and optical properties of the aerosols, and the quantitative parameterization of aerosol scattering and absorption. These factors could potentially modulate the relationship between SWA and PW in models, and thus  $dSWA/dPW$  may not completely reflect the physics of water vapour absorption. One possible example of this is with the GFDL models. All three GFDL models (GFDL-CM3, GFDL-ESM2G, GFDL-ESM2M) use the same shortwave parameterization for water vapour absorption and scattering by aerosols<sup>34–37</sup>; yet the CM3 absorbs more solar radiation in moist conditions, leading to larger  $dSWA/dPW$  (Extended Data Fig. 5). The CM3 implements an interactive aerosol scheme with different aerosol optical properties from the ESM models (which prescribe aerosols), resulting in enhanced and more realistic downward clear-sky surface shortwave flux due to reduced aerosol direct effects<sup>34</sup>. We speculate that the reduced aerosol direct effects may lead to more solar radiation available for absorption by lower tropospheric water vapour, and thus larger SWA in moist conditions in the CM3. The extent that differences in aerosol concentrations/properties affect the variability in  $dSWA/dPW$  among other models is not known without a more rigorous and controlled investigation. However, we suspect these effects are generally small, as (1) the GFDL models represent the only instance in which  $dSWA/dPW$  substantially differs among models with similar parameterizations of solar absorption by water vapour (Fig. 4, Extended Data Fig. 6), and (2)  $dSWA/dPW$  computed using different forcing scenarios (for example, historical and RCP8.5, which potentially exhibit large variability in aerosols) are very similar to those computed with the piControl (not shown).



Stratospheric ozone is another strong absorber of solar radiation. Ozone may influence dSWA/dPW if its concentration varies systematically with atmospheric moisture. Although column-integrated ozone tends to decrease with increasing PW in 16 models that have available ozone output, there is no cross-model correlation between the sensitivity of ozone concentration to varying PW and dSWA/dPW (not shown). This suggests the covariability of ozone and PW does not systematically affect dSWA/dPW. Furthermore, dSWA/dPW is similar to that in Extended Data Fig. 5 when it is computed from a smaller sample of locations and months exhibiting little variability in column-integrated ozone (not shown). This strengthens the case that ozone is not significantly affecting model variability in dSWA/dPW. However, the parameterization of ozone absorption probably influences the mean position of the SWA versus PW curve for each model (that is, average SWA over the range of PW analysed). One possible example of this is with the INM-CM4 and MRI-CGCM3 models. Both models use the same shortwave parameterization for water vapour absorption<sup>38</sup>, but have different treatments of solar absorption by ozone<sup>39,40</sup>. We speculate that this partly explains the similar slopes but different vertical placement of the SWA versus PW curves for these models (Extended Data Fig. 5).

As discussed above, dSWA/dPW values are nearly invariant to the simulation (for example, piControl, abrupt4 × CO<sub>2</sub>, historical, RCP8.5) from which they are calculated in models (not shown). They are also very similar when computed from a subset of years as short as that used from observations (2001–2009). Furthermore, dSWA/dPW computed from various percentiles of the SWA distribution within each PW bin (ranging from the 10th to 90th percentile) are similar to those computed with the SWA bin mean (not shown). These findings further demonstrate that the methodology robustly quantifies the dependence of SWA on atmospheric moisture.

**Observations.** The CERES-EBAF data set provides clear-sky radiative fluxes at the TOA and surface on a grid comparable to climate models and has global coverage. TOA fluxes are based on satellite measurements, and surface fluxes are generated with a radiative transfer model and are constrained by the TOA fluxes<sup>22</sup>. Although the surface fluxes are model produced, they are computed with a radiative transfer algorithm that is arguably more advanced and physically based than that used in most climate models. In particular, the radiation code for solar absorption employs the formal correlated-*k*-distribution framework with absorption coefficients (*k* values) being determined directly from detailed line-by-line (LBL)-generated *k* distributions<sup>41,42</sup>. This approach is arguably superior to that in most CMIP5 models, in which *k* values are in many cases determined with non-physical mathematical optimization procedures (for example, refs 36 and 43). In addition, the treatment of pressure–temperature–concentration dependence of *k* values in the CERES-EBAF scheme is more physical and higher in resolution than most CMIP5 models<sup>41</sup>. The final parameterization describing water vapour absorption in CERES-EBAF also has many mathematical terms (>50) approximating shortwave transmission<sup>44</sup>. CERES-EBAF surface fluxes are in good agreement with point observations<sup>22</sup> and the radiation scheme used to generate surface fluxes performs well when compared with recent LBL calculations<sup>24</sup>.

To compute CERES-EBAF-derived dSWA/dPW, water vapour data are taken from three sources: (1) the Special Sensor Microwave Imager (SSM/I)<sup>45</sup>; (2) a product developed by Remote Sensing Systems (RSS) that combines measurements from various instruments, including SSM/I, the Special Sensor Microwave Imager Sounder (SSMIS), the Advanced Microwave Scanning Radiometer (AMSR-E), and the WindSat Polarimetric Radiometer<sup>46</sup>; and (3) the Television Infrared Observation Satellite Operational Vertical Sounder (TOVS)<sup>47</sup>. ISCCP-FD-derived dSWA/dPW is based on TOVS water vapour, which was used in the development of the flux data set<sup>27</sup>.

The uncertainty range of CERES-EBAF-derived dSWA/dPW shown in the figures in this Letter (for example, Figs 3 and 4) reflects statistical uncertainty in the computation of dSWA/dPW and uncertainty due to the use of different PW data sets (see Fig. 3 legend). Inherent uncertainties in CERES-EBAF fluxes, including possible measurement uncertainty, errors in the radiative transfer scheme, and uncertainty in the methodology that generates clear-sky fluxes<sup>22</sup>, are difficult to quantify and are not included. Potential uncertainty in the individual PW measurements are also not accounted for. Thus, the effective uncertainty of observed dSWA/dPW is probably larger than indicated in the figures. Nonetheless, the conclusion that most CMIP5 models underestimate dSWA/dPW is robust, as it is strongly supported by a recent study in which radiation schemes are evaluated against high quality LBL calculations<sup>25</sup>.

**Additional discussion of shortwave radiative transfer schemes.** The severe underestimation of dSWA/dPW by the GISS models stands out in Fig. 4. In these models, solar absorption by water vapour is parameterized with a pseudo-*k*-distribution approach consisting of 15 mathematical terms<sup>48–52</sup>. Some of the other CMIP5 models in our analysis, including those with the largest dSWA/dPW, use as many terms in their parameterizations (Fig. 4). Thus the poor performance

of the GISS parameterization is not simply the result of the number of computations employed to approximate shortwave transmission. Rather, the finer details of how the parameters of the analytical expressions (for example, pseudo absorption coefficients and weights for terms) are developed and the quality of the reference calculations from which the parameterizations were originally based, are probably important, among other characteristics. The GISS parameterization employed in CMIP5 was developed from a combination of old and relatively new methods<sup>50–52</sup>. The resulting analytical expressions, which combine pressure–temperature–spectral absorption dependency, are known to underestimate solar absorption in moist atmospheres based on comparison with modern LBL calculations<sup>24,25</sup>. Updates were made to the GISS radiation scheme since CMIP5, and these will probably result in significant improvements in solar absorption and dSWA/dPW in future generations of the GISS model<sup>25</sup>.

Apart from the GISS models, the number of mathematical terms employed in shortwave parameterizations appear to exert a general influence on parameterization performance (Fig. 4). A specific example is with models developed at IPSL and CNRM. In these models, water vapour absorption is parameterized with an algorithm originally developed in 1980 and later modified for use in the operational European Center for Medium Range Weather Forecasts (ECMWF) model<sup>53,54</sup>. It consists of a few shortwave bands, within which Padé Approximants represent gaseous absorption by water vapour<sup>53</sup>. In the IPSL models, only two shortwave bands are used<sup>54</sup>. These models clearly underestimate mean SWA over the range of PW analysed and dSWA/dPW is smaller than that of all models except GISS (Extended Data Fig. 5). By contrast, the CNRM scheme employs a total of 6 shortwave bands<sup>28</sup>. Mean SWA is considerably larger and more realistic in these models compared to IPSL and dSWA/dPW is marginally improved as well (Extended Data Fig. 5). Thus the number of spectral bands and corresponding computations can have a large impact on the realism of solar absorption.

A comparison of the models that implement a 7-band parameterization originally developed in ref. 38 (BCC-CSM1.1, BCC-CSM1.1(m), CCSM4, INM-CM4, MRI-CGCM3, NorESM1-M) sheds light on specific characteristics of parameterizations of gaseous absorption, other than number of mathematical terms, that appear important for SWA. The original parameterization for water vapour absorption in ref. 38 consists of a 7-term pseudo-*k*-distribution summation with absorption coefficients and weights determined by fits to empirical and LBL calculations, respectively<sup>38,50</sup>. It is employed by the INM-CM4 and MRI-CGCM3 models (Extended Data Fig. 6). The parameterization was later modified to account for additional near-infrared water vapour absorption based on updated spectroscopic data and continuum absorption in the shortwave, which resulted in refitting the 7 parameterized absorption coefficients<sup>55</sup>. The updated parameterization is employed in the BCC-CSM1.1, BCC-CSM1.1(m), CCSM4, and NorESM1-M models. The models using the updated parameterization exhibit improved dSWA/dPW by a small but non-negligible amount (Extended Data Fig. 6). Consideration of weak water vapour absorption lines and continuum absorption are therefore somewhat important for accurate simulation of changes in solar absorption in a warming climate. This is consistent with the findings presented in refs 55 and 56. Note that most models with larger and more realistic dSWA/dPW tend to account for continuum absorption in their parameterizations (Extended Data Fig. 6).

Other details of the CMIP5 shortwave parameterization schemes, including the treatment of scattering by aerosols and molecules and of overlapping absorption by multiple gaseous species, have not been thoroughly investigated here. They too may influence the intermodel spread in dSWA/dPW. Even if these details are indeed influencing dSWA/dPW, it would not change the conclusion that shortwave parameterizations in general are important for the spread in simulated hydrologic cycle intensification.

**Constraining late twenty-first century precipitation changes.** We exploit the strong model relationships among dSWA/dPW, dSWA/dT, and *L<sub>v</sub>*dP/dT (Extended Data Fig. 7a, b) to compute a hypothetical change in *L<sub>v</sub>*P that may occur at the end of the twenty-first century under realistic climate forcing if the true temperature-mediated SWA response to CO<sub>2</sub> forcing, dSWA/dT, were perfectly known. The ‘true’ dSWA/dT is approximated from the model relationship between dSWA/dT and dSWA/dPW, using the observed value of dSWA/dPW based on CERES-EBAF (Extended Data Fig. 7a). The resulting true value of dSWA/dT is then used with the model relationship between the temperature-mediated *L<sub>v</sub>*P response, *L<sub>v</sub>*dP/dT, and dSWA/dT to estimate a ‘bias’ in *L<sub>v</sub>*dP/dT that originates from a bias in dSWA/dT (Extended Data Fig. 7b). The bias for each model is then removed from the predicted precipitation change at the end of the twenty-first century in the RCP8.5 scenario relative to the piControl according to:

$$\frac{L_v \Delta P}{\Delta T}_{\text{cons}} = \frac{L_v \Delta P_{\text{RCP8.5}} - (\text{bias} \times \Delta T_{\text{RCP8.5}})}{\Delta T_{\text{RCP8.5}}} \quad (2)$$

where  $\frac{L_v \Delta P}{\Delta T_{\text{cons}}}$  is the constrained total change in  $L_v P$  normalized by surface warming with the bias removed,  $L_v \Delta P_{\text{RCP8.5}}$  is the total late twenty-first century  $L_v P$  change (mean of years 2081–2100 in RCP8.5 minus mean of years 131–150 in piControl), and  $\Delta T_{\text{RCP8.5}}$  is the late twenty-first century 2-m air temperature change computed similarly to  $L_v \Delta P_{\text{RCP8.5}}$ .

The above procedure makes several assumptions, including: (1) the middle of the range in CERES-EBAF-computed dSWA/dPW is most representative of the real atmosphere; (2) the best 'true' value of dSWA/dT (Extended Data Fig. 7a, blue star) and  $L_v dP/dT$  (Extended Data Fig. 7b, black horizontal line) occurs at the linear regression line on the cross-model scatterplots; and (3)  $L_v dP/dT$  contributes linearly (with  $\Delta T$ ) to the total late twenty-first century change in  $L_v P$  computed from RCP8.5, as depicted in equation (2).

Removing the bias in  $L_v dP/dT$  due to a bias in dSWA/dT reduces the model spread in predicted precipitation change per unit warming at the end of the twenty-first century by 37%, and reduces the ensemble mean increase by 38% (Extended Data Fig. 7c). Even if we do not normalize by differences in surface warming  $\Delta T$ , which is the main driver of the spread in total precipitation change  $L_v \Delta P$ , a discernible reduction in spread by 27% and ensemble mean increase by 25% can be achieved (Extended Data Fig. 7d). The spread reduction is substantial, considering the numerous factors in addition to the temperature-mediated SWA response to CO<sub>2</sub> forcing potentially contributing to model scatter in RCP8.5 projections. These include temperature-mediated responses of other energy budget components (Extended Data Fig. 2), greenhouse-gas forcing other than CO<sub>2</sub>, aerosols, and the rapid  $L_v P$  adjustments to all forcings. For instance, ref. 4 demonstrated the potent role of black carbon forcing for CMIP3 spread in simulated global precipitation change under a realistic climate change scenario. That we obtain a 37% reduction in  $L_v \Delta P/\Delta T$  under RCP8.5 by only constraining the temperature-mediated component of SWA change under pure CO<sub>2</sub> forcing (and only a somewhat larger reduction by 45% when repeating the same exercise with the quadrupled CO<sub>2</sub> runs, not shown) suggests that the role of black carbon forcing on the spread may be less potent in CMIP5 than CMIP3. This is an interesting possibility worthy of further analysis.

How may uncertainty in late twenty-first century precipitation change be reduced further? As discussed above, a realistic climate change scenario includes greenhouse-gas forcing (from CO<sub>2</sub> and other gases), aerosol forcing, and rapid adjustments to these forcings. A better understanding of all these factors is therefore critical, including understanding of the rapid adjustment to CO<sub>2</sub> forcing. This factor has a non-negligible spread (Extended Data Fig. 3b) and is not strongly correlated with the corresponding intermodel variations in temperature-mediated response ( $r = -0.23$ , not shown). Additionally, the spread in the temperature-mediated  $L_v P$  response to CO<sub>2</sub> forcing is not only driven by the SWA component, as indicated by residual scatter in Fig. 2b. Extended Data Fig. 2 shows that the net atmospheric longwave cooling response, dLWC/dT, also has a large spread that is correlated with  $L_v dP/dT$ . dSWA/dT and dLWC/dT are not correlated with each other ( $|r| < 0.1$ , not shown), suggesting that dLWC/dT is another independent source of spread that demands better understanding. dLWC/dT is only correlated with  $L_v dP/dT$  for all-sky (Extended Data Fig. 2), implying that clouds may play an important role in the intermodel relationship. This is different from the case of dSWA/dT, in which clear-sky absorption by water vapour is critical.

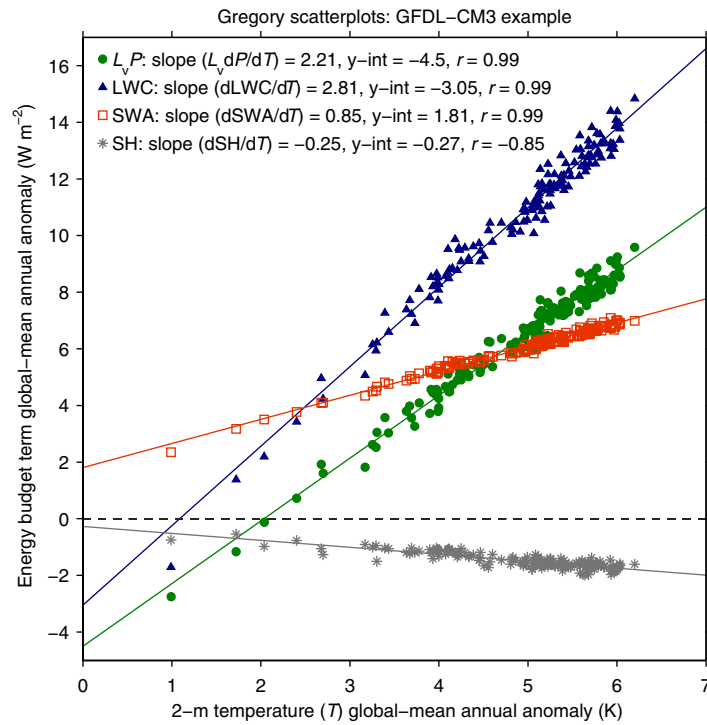
**Sample size.** No statistical methods were used to predetermine sample size.

**Code availability.** Any codes used in the analysis in this paper and in the production of figures can be made available upon request. Please contact A.M.D. (adeangelis@ucla.edu).

30. Andrews, T., Forster, P. M. & Gregory, J. M. A surface energy perspective on climate change. *J. Clim.* **22**, 2557–2570 (2009).
31. Lu, J. & Cai, M. Stabilization of the atmospheric boundary layer and the muted global hydrological cycle response to global warming. *J. Hydrometeorol.* **10**, 347–352 (2009).
32. Allan, R. P. *et al.* Physically consistent responses of the global atmospheric hydrological cycle in models and observations. *Surv. Geophys.* **35**, 533–552 (2014).
33. Devore, J. L. in *Probability and Statistics for Engineering and the Sciences* 7th edn (eds Crockett, C., Gershman, B. & Broyer, R.) 12.446–12.499 (Duxbury Press, Belmont, California, 2008).
34. Donner, L. J. *et al.* The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3. *J. Clim.* **24**, 3484–3519 (2011).
35. Dunne, J. P. *et al.* GFDL's ESM2 global coupled climate-carbon Earth system models. Part I: Physical formulation and baseline simulation characteristics. *J. Clim.* **25**, 6646–6665 (2012).
36. Freidenreich, S. M. & Ramaswamy, V. A new multiple-band solar radiative parameterization for general circulation models. *J. Geophys. Res.* **104**, 31389–31409 (1999).
37. The GFDL Global Atmospheric Model Development Team. The new GFDL global atmosphere and land model AM2–LM2: evaluation with prescribed SST simulations. *J. Clim.* **17**, 4641–4673 (2004).
38. Briegleb, B. P. Delta-eddington approximation for solar radiation in the NCAR Community Climate Model. *J. Geophys. Res.* **97**(D7), 7603–7612 (1992).
39. Galin, V. Y. Parametrization of radiative processes in the DNM atmospheric model. *Izvest. Atmos. Ocean. Phys.* **34**, 339–347 (1998).
40. Yukimoto, S. *et al.* Meteorological Research Institute-Earth System Model Version 1 (MRI-ESM1) — Model description. Tech. Rep. No. 64 (Meteorological Research Institute, Tsukuba-city, Ibaraki, Japan, 2011).
41. Kato, S., Ackerman, T. P., Mather, J. H. & Clothiaux, E. E. The k-distribution method and correlated-k approximation for a shortwave radiative transfer model. *J. Quant. Spectrosc. Radiat. Trans.* **62**, 109–121 (1999).
42. Kato, S. *et al.* Improvements of top-of-atmosphere and surface irradiance computations with CALIPSO-, CloudSat-, and MODIS-derived cloud and aerosol properties. *J. Geophys. Res.* **116**, D19209 (2011).
43. Sekiguchi, M. & Nakajima, T. A k-distribution-based radiation code and its computational optimization for an atmospheric general circulation model. *J. Quant. Spectrosc. Radiat. Trans.* **109**, 2779–2793 (2008).
44. Rose, F. *et al.* CERES Proto-Edition 3 radiative transfer: tests and radiative closure over surface validation sites. In *12th Conference on Atmospheric Radiation* <https://ams.confex.com/ams/pdfpapers/112358.pdf> (2006).
45. Ferraro, R. R., Weng, F., Grody, N. C. & Basist, A. An eight-year (1987–1994) time series of rainfall, clouds, water vapor, snow cover, and sea ice derived from SSM/I measurements. *Bull. Am. Meteorol. Soc.* **77**, 891–905 (1996).
46. Remote Sensing Systems. Monthly mean total precipitable water data set on a 1 degree grid made from Remote Sensing Systems Version-7 microwave radiometer data. <http://www.remss.com/measurements/atmospheric-water-vapor/tpw-1-deg-product> (2012, updated July 2014, accessed August 2014).
47. Kidwell, K. *NOAA Polar Orbiter Data Products Users Guide* (National Environmental Satellite Data and Information Service, Silver Spring, Maryland, 1995).
48. Schmidt, G. A. *et al.* Present-day atmospheric simulations using GISS ModelE: comparison to in situ, satellite, and reanalysis data. *J. Clim.* **19**, 153–192 (2006).
49. Schmidt, G. A. *et al.* Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive. *J. Adv. Model. Earth Syst.* **6**, 141–184 (2014).
50. Lacis, A. A. & Hansen, J. E. A parameterization for the absorption of solar radiation in the Earth's atmosphere. *J. Atmos. Sci.* **31**, 118–133 (1974).
51. Hansen, J. *et al.* Efficient three-dimensional global models for climate studies: Models I and II. *Mon. Weath. Rev.* **111**, 609–662 (1983).
52. Lacis, A. A. & Oinas, V. A description of the correlated k distribution method for modeling nongray gaseous absorption, thermal emission, and multiple scattering in vertically inhomogeneous atmospheres. *J. Geophys. Res.* **96**(D5), 9027–9063 (1991).
53. Fouquart, Y. & Bonnel, B. Computations of solar heating of the Earth's atmosphere: a new parameterization. *Contrib. Atmos. Phys.* **53**, 35–62 (1980).
54. Morcrette, J.-J. Radiation and cloud radiative properties in the European Centre for Medium Range Weather Forecasts forecasting system. *J. Geophys. Res.* **96**(D5), 9121–9132 (1991).
55. Collins, W. D., Lee-Taylor, J. M., Edwards, D. P. & Francis, G. L. Effects of increased near-infrared absorption by water vapor on the climate system. *J. Geophys. Res.* **111**, D18109 (2006).
56. Paynter, D. & Ramaswamy, V. Investigating the impact of the shortwave water vapor continuum upon climate simulations using GFDL global models. *J. Geophys. Res.* **119**, 10720–10737 (2014).
57. Hewitt, H. T. *et al.* Design and implementation of the infrastructure of HadGEM3: the next-generation Met Office climate modelling system. *Geosci. Model Dev.* **4**, 223–253 (2011).
58. Bi, D. *et al.* The ACCESS coupled model: description, control climate and evaluation. *Austral. Meteorol. Oceanogr. J.* **63**, 41–64 (2013).
59. Edwards, J. M. & Slingo, A. Studies with a flexible new radiation code. I: Choosing a configuration for a large-scale model. *Q. J. R. Meteorol. Soc.* **122**, 689–719 (1996).
60. Cusack, S., Edwards, J. M. & Crowther, J. M. Investigating k distribution methods for parameterizing gaseous absorption in the Hadley Centre climate model. *J. Geophys. Res.* **104**, 2051–2057 (1999).
61. Wu, T. *et al.* The Beijing Climate Center atmospheric general circulation model: description and its performance for the present-day climate. *Clim. Dyn.* **34**, 123–147 (2010).
62. Collins, W. D. *et al.* Description of the NCAR Community Atmosphere Model (CAM 3.0). Tech. Rep. NCAR/TN-464+STR (National Center for Atmospheric Research, Boulder, Colorado, 2004).
63. Chylek, P., Li, J., Dubey, M. K., Wang, M. & Lesins, G. Observed and model simulated 20<sup>th</sup> century Arctic temperature variability: Canadian Earth System Model CanESM2. *Atmos. Chem. Phys. Discuss.* **11**, 22893–22907 (2011).
64. Li, J. & Barker, H. W. A radiation algorithm with correlated-k distribution. Part I: Local thermal equilibrium. *J. Atmos. Sci.* **62**, 286–309 (2005).
65. Gent, P. R. *et al.* The Community Climate System Model Version 4. *J. Clim.* **24**, 4973–4991 (2011).
66. Neale, R. B. *et al.* Description of the NCAR Community Atmosphere Model (CAM 4.0). Tech. Rep. NCAR/TN-485+STR (National Center for Atmospheric Research, Boulder, Colorado, 2010).

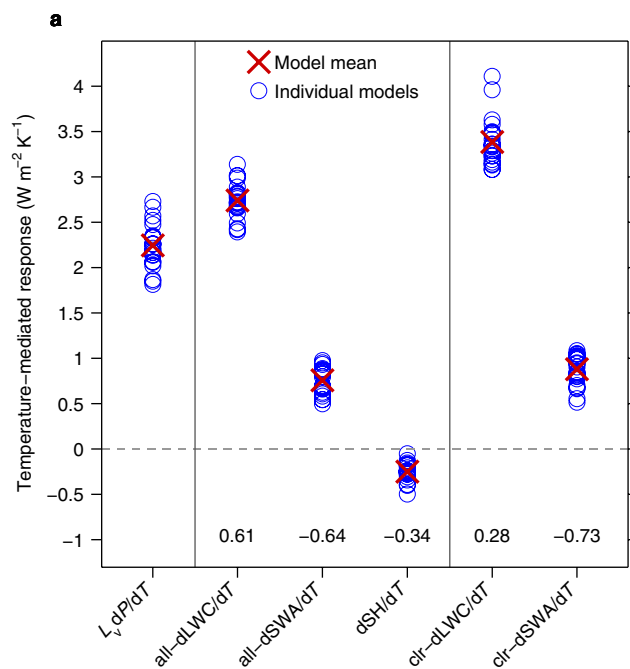
67. Martin, G. M. *et al.* The physical properties of the atmosphere in the new Hadley Centre Global Environmental Model (HadGEM1). Part I: Model description and global climatology. *J. Clim.* **19**, 1274–1301 (2006).
68. Collins, W. J. *et al.* Development and evaluation of an Earth-System model — HadGEM2. *Geosci. Model. Dev.* **4**, 1051–1075 (2011).
69. Volodin, E. M., Dianskii, N. A. & Gusev, A. V. Simulating present day climate with the INMCM4.0 coupled model of the atmospheric and oceanic general circulations. *Izvest. Atmos. Ocean. Phys.* **46**, 414–431 (2010).
70. Dufresne, J.-L. *et al.* Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Clim. Dyn.* **40**, 2123–2165 (2013).
71. Hourdin, F. *et al.* LMDZ5B: the atmospheric component of the IPSL climate model with revisited parameterizations for clouds and convection. *Clim. Dyn.* **40**, 2193–2222 (2013).
72. Watanabe, S. *et al.* MIROC-ESM 2010: model description and basic results of CMIP5-20c3m experiments. *Geosci. Model. Dev.* **4**, 845–872 (2011).
73. Watanabe, M. *et al.* Improved climate simulation by MIROC5: mean states, variability, and climate sensitivity. *J. Clim.* **23**, 6312–6335 (2010).
74. Jungclaus, J. H. *et al.* Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model. *J. Adv. Model. Earth Syst.* **5**, 422–446 (2013).
75. Stevens, B. *et al.* Atmospheric component of the MPI-M Earth system model: ECHAM6. *J. Adv. Model. Earth Syst.* **5**, 146–172 (2013).
76. Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J. & Clough, S. A. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.* **102**(D14), 16663–16682 (1997).
77. Mlawer, E. J. & Clough, S. A. On the extension of rapid radiative transfer model to the shortwave region. In *Proceedings of the 6th Atmospheric Radiation Measurement Science Team Meeting* [https://www.arm.gov/publications/proceedings/conf06/extended\\_abs/mlawer\\_ej.pdf](https://www.arm.gov/publications/proceedings/conf06/extended_abs/mlawer_ej.pdf) (US Department of Energy, CONF-9603149, 1997).
78. Iacono, M. J. *et al.* Radiative forcing by long-lived greenhouse gases: calculations with the AER radiative transfer models. *J. Geophys. Res.* **113**, D13103 (2008).
79. Yukimoto, S. *et al.* A new global climate model of the Meteorological Research Institute: MRI-CGCM3. *J. Meteorol. Soc. Jpn* **90A**, 23–64 (2012).
80. Bentsen, M. *et al.* The Norwegian Earth System Model, NorESM1-M—Part 1: Description and basic evaluation of the physical climate. *Geosci. Model. Dev.* **6**, 687–720 (2013).



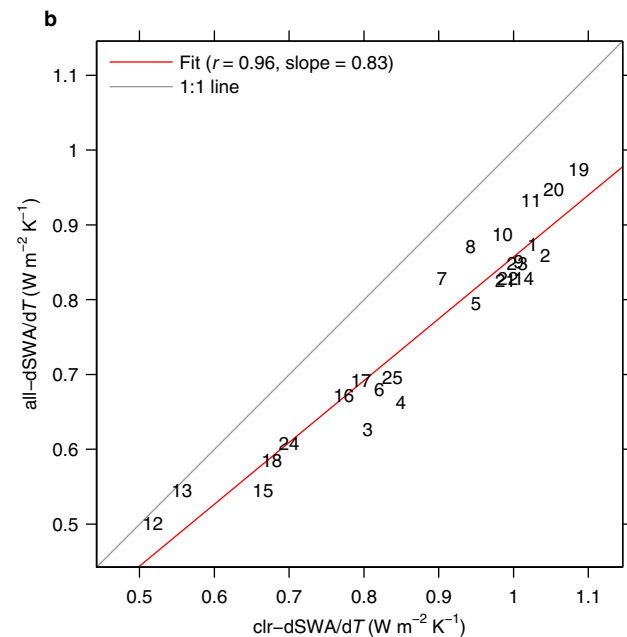


**Extended Data Figure 1 | Demonstration of the Gregory method for the GFDL-CM3.** Global-mean annual anomalies (abrupt4 × CO<sub>2</sub>—piControl; see Methods) in atmospheric energy budget terms (latent heat release from precipitation ( $L_vP$ ), net longwave cooling (LWC), shortwave absorption (SWA), and sensible heating (SH)) are regressed against those in 2-m air

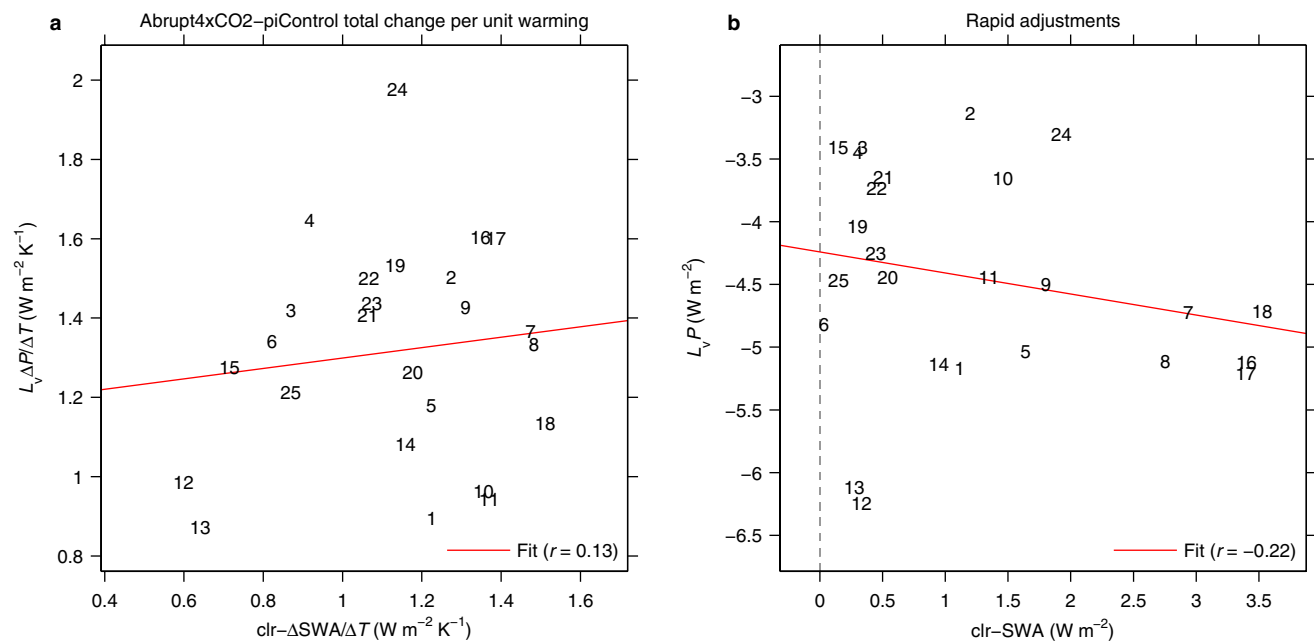
temperature (see Methods). For  $L_vP$ , precipitation anomalies are multiplied by the latent heat of vaporization,  $L_v$ . Radiative terms are computed with all-sky fluxes. The statistics of the linear regression (slope (temperature-mediated response in  $W m^{-2} K^{-1}$ ), y-intercept (rapid adjustment in  $W m^{-2}$ ), and correlation coefficient,  $r$ ) are displayed in the key.



**Extended Data Figure 2 | Summary of model spread in temperature-mediated responses.** **a**, The temperature-mediated response of each atmospheric energy budget term (equation (1) in main text) is shown for each model as blue circles and the model mean as a red cross. Responses of the radiative terms (dLWC/dT and dSWA/dT) computed with all-sky



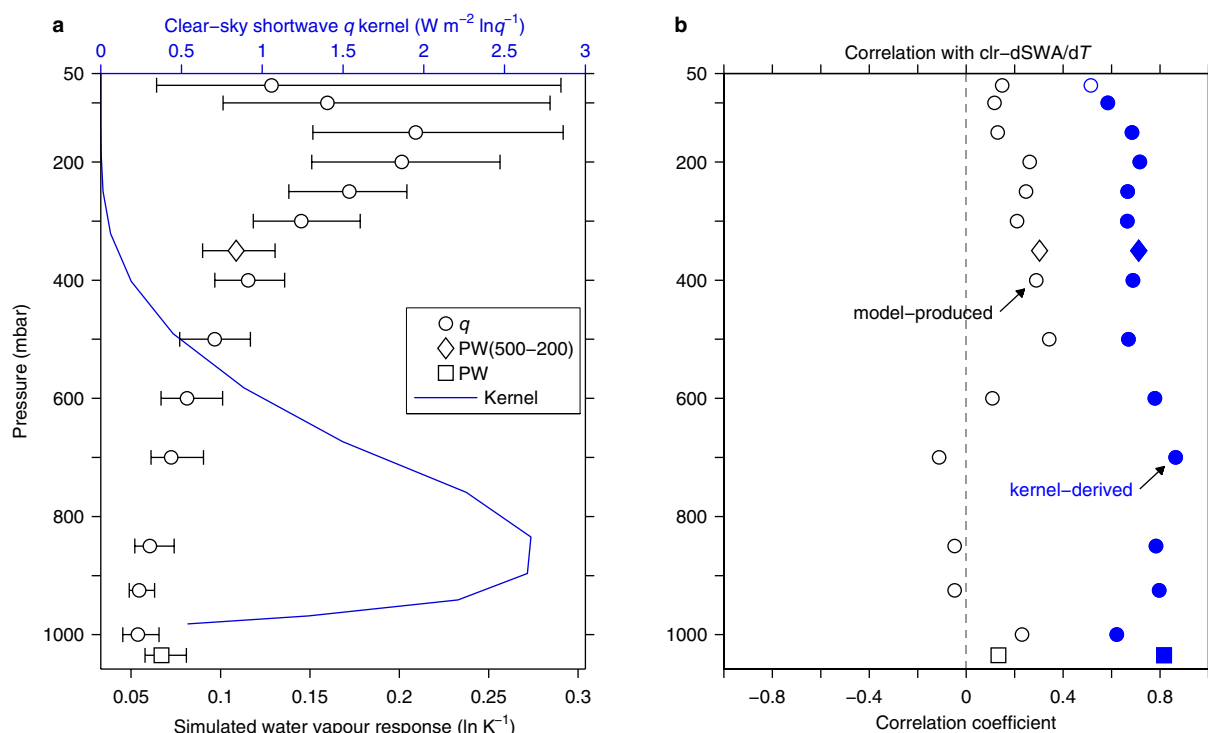
fluxes or with clear-sky fluxes are denoted with all- or clr-, respectively. The numbers above the abscissa are the cross-model correlations between  $L_v dP/dT$  and each other temperature-mediated response. **b**, Scatterplot of the all-sky versus clear-sky temperature-mediated SWA response and corresponding linear fit.



**Extended Data Figure 3 | Total changes and rapid adjustments for  $L_v P$  and SWA.** **a**, The total change in  $L_v P$  per unit warming (mean over years 131–150 of the abrupt4 × CO<sub>2</sub> simulation minus the corresponding mean of the piControl simulation, normalized by 2-m air temperature change,

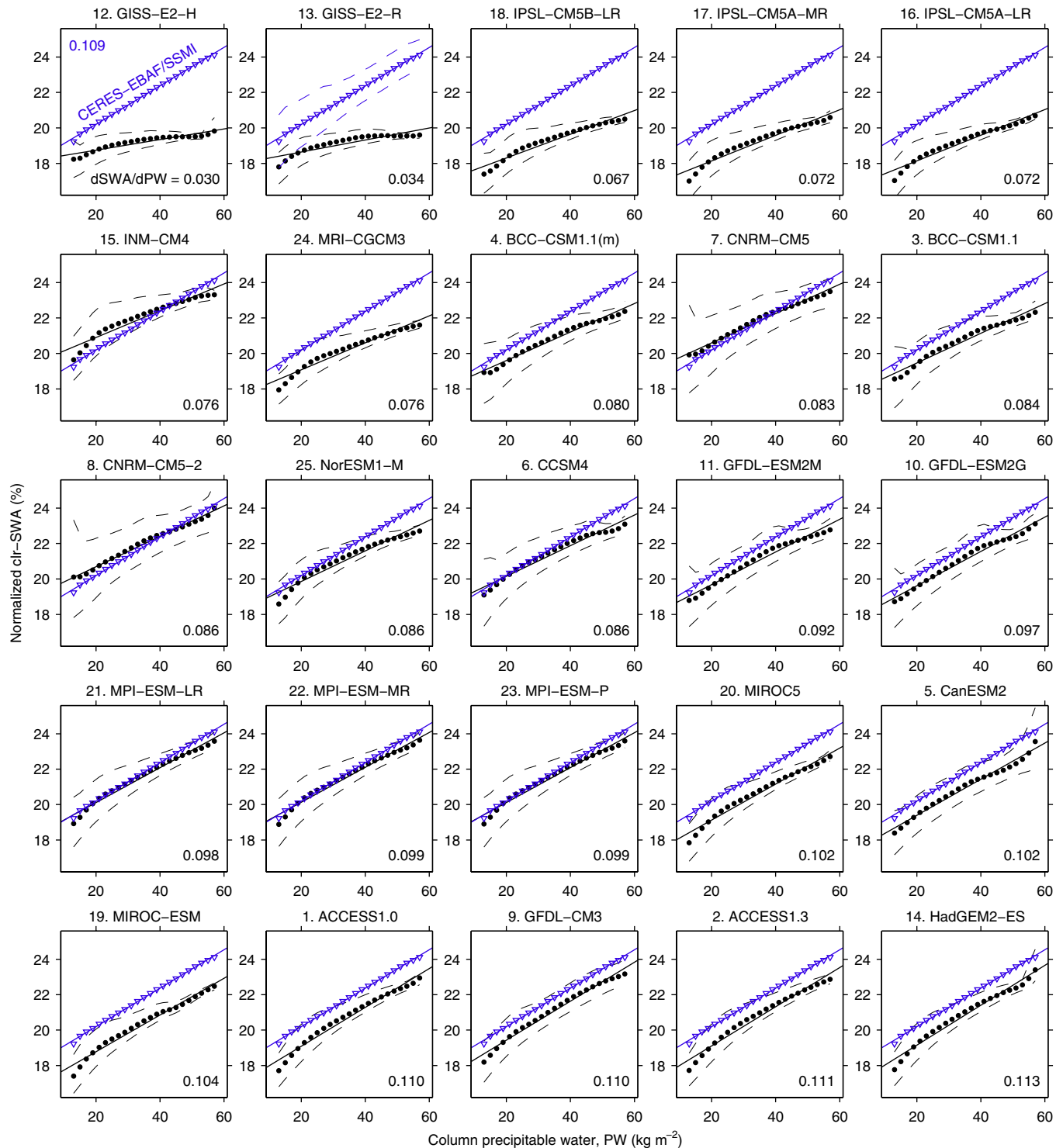
$L_v P / \Delta T$ ) versus total change in clear-sky SWA per unit warming ( $\text{clr-}\Delta\text{SWA} / \Delta T$ ). **b**, The rapid adjustment of  $L_v P$  versus rapid adjustment of clear-sky ( $\text{clr-}$ ) SWA. A linear fit is shown in both cases.





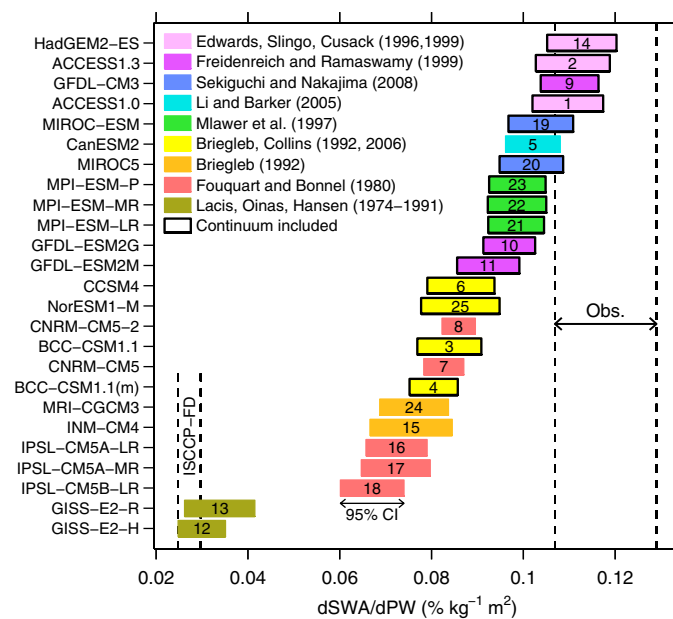
**Extended Data Figure 4 | Contributions of water vapour change to model spread in temperature-mediated SWA response.** **a**, The Gregory method (Methods) is applied to anomalies of globally averaged specific humidity ( $q$ ) at standard atmospheric levels, total column precipitable water (PW), and upper tropospheric precipitable water (PW(500–200), computed by vertically integrating  $q$  between 500 and 200 mbar) to quantify the temperature-mediated response of atmospheric water vapour for each model. The natural log was taken before computing annual anomalies. For each quantity, the symbols (circle, diamond,

square) represent the model mean and the whiskers represent the full model spread. The globally averaged annual-mean clear-sky shortwave atmospheric  $q$  kernel (Methods) is overlaid (blue curve). **b**, The cross-model correlation between the responses of water vapour in **a** and the temperature-mediated clear-sky SWA response ( $\text{clr-dSWA/dT}$ ) computed with model-produced fluxes (black) or radiative kernels (blue). Filled symbols are statistically significant at the 5% level based on a two-tailed  $t$ -test<sup>33</sup>, with degrees of freedom corresponding to the number of participating modelling institutions (14) within the 25 model ensemble.



**Extended Data Figure 5 | The SWA sensitivity curve for each model.** Shown are normalized bin-mean SWA versus PW and corresponding linear fit (as in Fig. 3a) for each model (black dots/line); models are sorted from (top left) smallest  $dSWA/dPW$  ( $\% \text{ kg}^{-1} \text{ m}^2$ ), printed on every panel) to (bottom right) largest  $dSWA/dPW$ . Dashed lines depict the 10th–90th percentile spread of SWA within each PW bin, demonstrating the tight

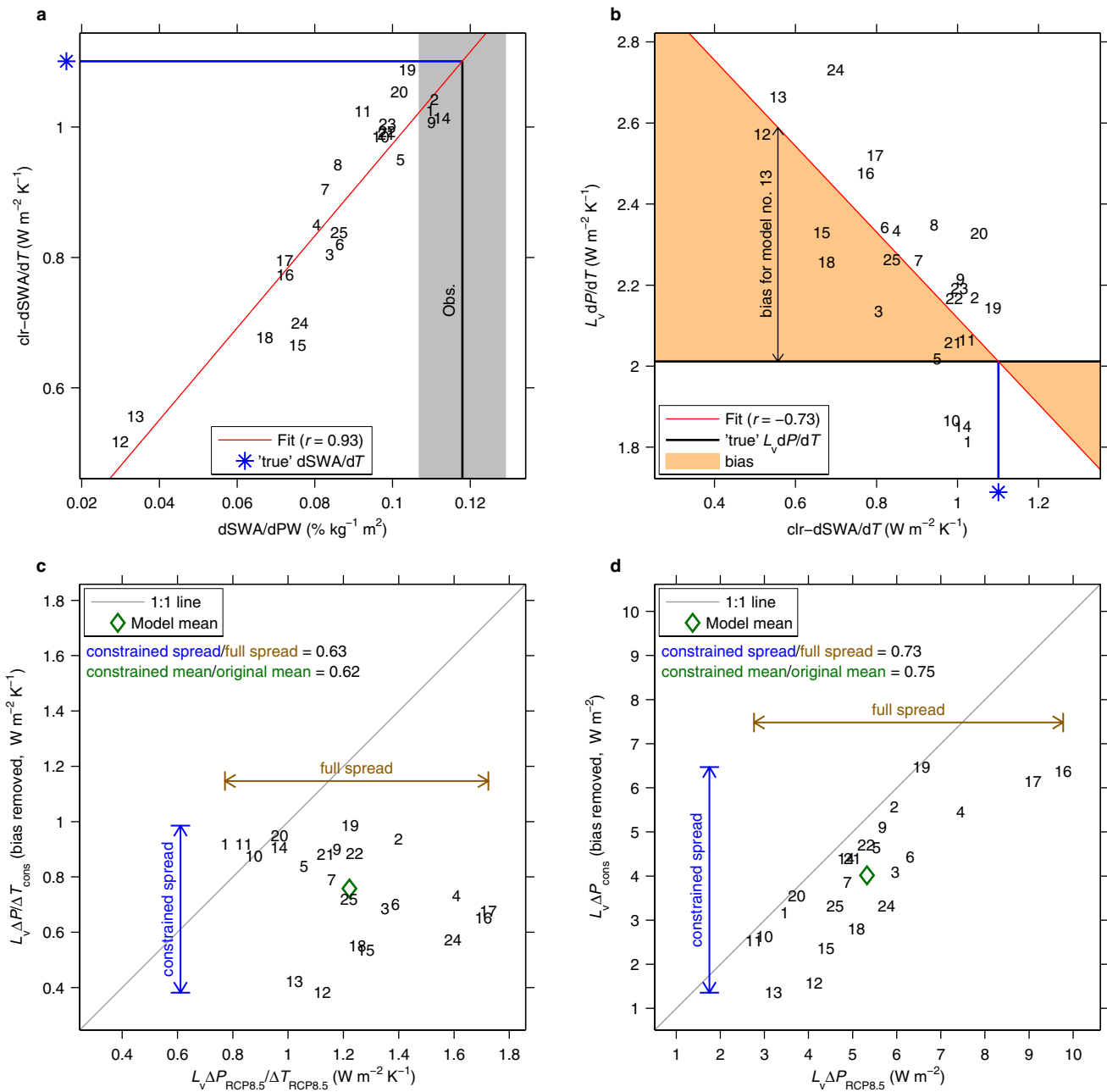
constraint PW places on SWA. Numbers next to model names are those from Extended Data Table 1. On every panel, the SWA versus PW curve and linear fit based on CERES-EBAF fluxes and SSM/I water vapour are also shown (blue triangles/line); the 10th–90th percentile spread is shown only on the second panel for visual clarity.



**Extended Data Figure 6 | Methodology for parameterizing absorption of solar radiation by water vapour.** Shown is the relationship between dSWA/dPW and methodology used to parameterize solar absorption by water vapour in a cloud-free atmosphere, with colours for each model referring to different parameterization procedures as documented in the references listed in the key (see also Extended Data Table 1). (Boxes outlined in black indicate that water vapour continuum absorption in the shortwave is accounted for in the parameterization.)

Model numbers, printed on each box, are identified on the ordinate axis as in Fig. 4. The width of the horizontal shading for each model represents the 95% CI of the regression slope to the SWA versus PW curve, as in Fig. 4. Statistical uncertainty of dSWA/dPW derived from CERES-EBAF fluxes and three PW data sets (Obs.) and from ISCCP-FD is represented with vertical dashed lines, as described in Figs 3 and 4, respectively. (refs 36, 38, 43, 50–53, 55, 59–60, 64, 76).





**Extended Data Figure 7 | Constraining the spread in late twenty-first century precipitation change.** **a**, The relationship between temperature-mediated clear-sky SWA response ( $\text{clr-dSWA/dT}$ ) and SWA sensitivity to varying PW ( $\text{dSWA/dPW}$ ) (as in Fig. 3b), showing an estimate of the 'true'  $\text{dSWA/dT}$  (blue line/star), and how it is quantified. **b**, The temperature-mediated  $L_v P$  response ( $L_v dP/dT$ ) versus  $\text{dSWA/dT}$  (as in Fig. 2b), showing how the 'true'  $\text{dSWA/dT}$  in **a** is used to quantify a bias in  $L_v dP/dT$  originating from a bias in  $\text{dSWA/dT}$ ; the bias for an example

model (no. 13, GISS-E2-R) is displayed. **c**, The full (abscissa) versus constrained (with bias in **b** removed, ordinate) total change (indicated with  $\Delta$ ) in  $L_v P$  normalized by change in 2-m warming at the end of the twenty-first century under RCP8.5 (see Methods and equation (2) in Methods). **d**, As in **c** but for total  $L_v P$  change not normalized by warming. Model numbers are defined in Extended Data Table 1. Two models (no. 8, CNRM-CM5-2 and no. 23, MPI-ESP-P) are excluded from **c** and **d** owing to unavailable RCP8.5 output.

**Extended Data Table 1 | CMIP5 models analysed in this paper**

Model	Modelling Institution(s), Country	References	
		General	SW code
1. ACCESS1.0	Commonwealth Scientific and Industrial Research Organization and Bureau of Meteorology, Australia	[57], [58]	[59], [60]
2. ACCESS1.3			
3. BCC-CSM1.1	Beijing Climate Center, China Meteorological Administration, China ( <a href="http://forecast.bccsm.cma.gov.cn/web/channel-1.htm">http://forecast.bccsm.cma.gov.cn/web/channel-1.htm</a> )	[61]	[55], [62]
4. BCC-CSM1.1(m)			
5. CanESM2	Canadian Centre for Climate Modelling and Analysis, Canada	[63]	[64]
6. CCSM4	National Center for Atmospheric Research, USA	[65]	[55], [66]
7. CNRM-CM5	Centre National de Recherches Meteorologiques / Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique, France	[28]	[28], [53]
8. CNRM-CM5-2			
9. GFDL-CM3	NOAA Geophysical Fluid Dynamics Laboratory, USA	[34]	[36], [37]
10. GFDL-ESM2G		[35]	
11. GFDL-ESM2M			
12. GISS-E2-H	NASA Goddard Institute for Space Studies, USA	[48], [49]	[50]–[52]
13. GISS-E2-R			
14. HadGEM2-ES	Met Office Hadley Centre, UK	[67], [68]	[59], [60]
15. INM-CM4	Institute for Numerical Mathematics, Russia	[69]	[38], [39]
16. IPSL-CM5A-LR	Institut Pierre-Simon Laplace, France	[70], [71]	[53], [54]
17. IPSL-CM5A-MR			
18. IPSL-CM5B-LR			
19. MIROC-ESM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies, Japan	[72]	[43]
20. MIROC5		[73]	
21. MPI-ESM-LR	Max Planck Institute for Meteorology, Germany	[74], [75]	[76]–[78]
22. MPI-ESM-MR			
23. MPI-ESM-P			
24. MRI-CGCM3	Meteorological Research Institute, Japan	[79]	[38], [40]
25. NorESM1-M	Norwegian Climate Centre, Norway	[80]	[55], [66]

The numbers used to identify models throughout the paper are given in the first column. General references and references documenting the details of the shortwave parameterization schemes, particularly with regard to the treatment of gaseous absorption, are listed in the rightmost columns (refs 28, 34–40, 43, 48–55, 57–80). Additional information about the BCC-CSM1.1 models may be found at the webpage given under the modelling institution.

# Scale dependence of rock friction at high work rate

Futoshi Yamashita<sup>1</sup>, Eiichi Fukuyama<sup>1</sup>, Kazuo Mizoguchi<sup>1,2</sup>, Shigeru Takizawa<sup>1</sup>, Shiqing Xu<sup>1</sup> & Hironori Kawakata<sup>1,3</sup>

Determination of the frictional properties of rocks is crucial for an understanding of earthquake mechanics, because most earthquakes are caused by frictional sliding along faults. Prior studies using rotary shear apparatus<sup>1–13</sup> revealed a marked decrease in frictional strength, which can cause a large stress drop and strong shaking, with increasing slip rate and increasing work rate. (The mechanical work rate per unit area equals the product of the shear stress and the slip rate.) However, those important findings were obtained in experiments using rock specimens with dimensions of only several centimetres, which are much smaller than the dimensions of a natural fault (of the order of 1,000 metres). Here we use a large-scale biaxial friction apparatus with metre-sized rock specimens to investigate scale-dependent rock friction. The experiments show that rock friction in metre-sized rock specimens starts to decrease at a work rate that is one order of magnitude smaller than that in centimetre-sized rock specimens. Mechanical, visual and material observations suggest that slip-evolved stress heterogeneity on the fault accounts for the difference. On the basis of these observations, we propose that stress-concentrated areas exist in which frictional slip produces more wear materials (gouge) than in areas outside, resulting in further stress concentrations at these areas. Shear stress on the fault is primarily sustained by stress-concentrated areas that undergo a high work rate, so those areas should weaken rapidly and cause the macroscopic frictional strength to decrease abruptly. To verify this idea, we conducted numerical simulations assuming that local friction follows the frictional properties observed on centimetre-sized rock specimens. The simulations reproduced the macroscopic frictional properties observed on the metre-sized rock specimens. Given that localized stress concentrations commonly occur naturally, our results suggest that a natural fault may lose its strength faster than would be expected from the properties estimated from centimetre-sized rock samples.

Large fault slips, which can produce strong ground motion and a huge tsunami, are associated with a remarkable decrease in rock friction at seismic slip rates. Numerous laboratory experiments<sup>1–13</sup> have investigated the underlying mechanisms of reduction of friction (weakening) at high slip rate. Although various weakening mechanisms—such as gouge or powder lubrication<sup>8,9,11</sup>, silica gel lubrication<sup>2,4</sup> and melting<sup>1,6,10,13</sup>—have been suggested, all of them seem to be triggered by a rapid temperature rise caused by frictional work on the fault<sup>5</sup>. Therefore, the mechanical work rate per unit area (that is, the power density) seems likely to be a key parameter.

Another question then arises: are the frictional properties independent of length scale? It is known that rupture propagation on a fault is affected by the fault surface topography<sup>14</sup>, and thus should be dependent on scale. Presumably, steady-sliding rock friction is also affected by scale. However, it has not been possible to investigate such a dependence because it has been difficult to conduct a larger-scale friction experiment. Although some friction experiments using metre-sized rock specimens have been conducted<sup>15–18</sup>, the apparatuses used could not produce large slip distances with high slip rates.

We therefore conducted experiments with metre-sized rock specimens by using a second-generation large-scale biaxial friction

apparatus at the National Research Institute for Earth Science and Disaster Prevention (NIED) in Japan (Fig. 1a)<sup>19</sup>. We used a pair of quadrangular prismatic rock specimens made of Indian metagabbro, whose nominal contact area was 1.5 m long and 0.1 m wide. The contact surfaces were flattened to an undulation of within 10  $\mu\text{m}$  before the first experiment. We repeated several experiments using the same rock specimens under the conditions of constant loading rates up to  $3 \times 10^{-2} \text{ m s}^{-1}$  and with constant normal stresses up to 6.7 MPa (see Supplementary Information section 1 and Supplementary Table 2 for detailed methods and conditions, respectively).

After each experiment, we found indications of localized damage on the fault surface, including gouge and grooves (Fig. 1b–d). The grooves began with very sharp edges (Fig. 1c) and widened with continued slip. The grooves were filled with the gouge, which looked heavily comminuted and partly consolidated (Fig. 1d). Consolidated gouge within the grooves often swelled up and spilled into the surrounding fault surface. The fault surface without grooves became smoother and smoother with experiments, which was also confirmed by topography measurements (see sections 2 and 3 in Supplementary Information for detailed observation of the fault surface).

After each experiment, we removed and inspected all gouge, including from inside the grooves. Therefore, the groove areas were not in contact at the beginning of the next experiment. To investigate the distribution of damage, we took pictures of the fault surface just after each experiment as well as after the gouge collection. Detailed methods of the damage analysis are described in Supplementary Information section 4.

Figure 2 shows a clear dependence of rock friction in the metre-sized rock specimens on the work rate per unit area (or power density), which we term ‘work rate’ for simplicity hereafter, obtained from the present experiments. At a low work rate, the rock friction obeys Byerlee’s law<sup>20</sup>, but it falls dramatically at high work rates. Such work-rate dependence has already been reported, based on the data of rotary shear experiments<sup>5,12</sup>.

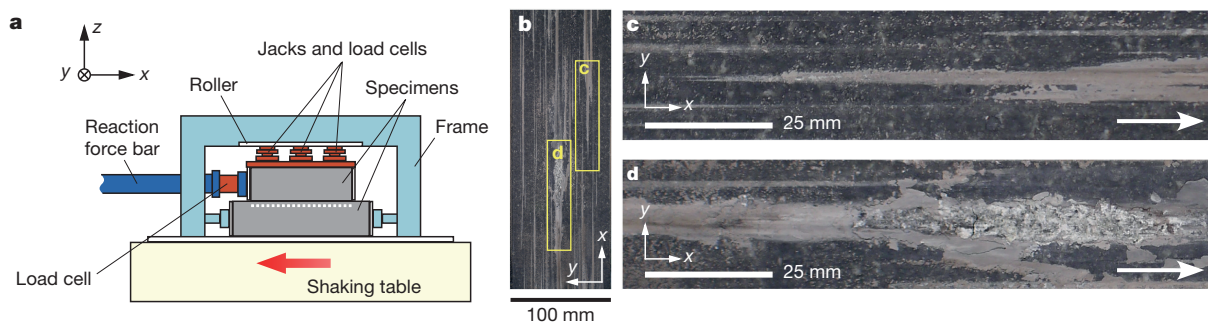
To directly investigate whether there is any size effect on the frictional properties, we also conducted rotary shear experiments with specimens of metagabbro and monzodiorite whose inner and outer diameters were 1.7 cm and 4.0 cm, respectively. The two rock types yield a similar frictional dependence on work rate, as shown in Fig. 2. The experimental data are listed in Supplementary Table 3. As shown in Fig. 2, the findings clearly demonstrate a scale dependence of rock friction, as follows. The friction in centimetre-sized rock specimens temporarily decreases at the intermediate work rate ( $\sim 5 \times 10^{-3} \text{ MJ m}^{-2} \text{ s}^{-1}$ ), increases with the increase in the work rate from  $10^{-2} \text{ MJ m}^{-2} \text{ s}^{-1}$  to  $10^{-1} \text{ MJ m}^{-2} \text{ s}^{-1}$  and suddenly falls at a work rate of  $10^{-1} \text{ MJ m}^{-2} \text{ s}^{-1}$ , whereas the friction in metre-sized rock specimens starts to decrease at a work rate that is one order of magnitude smaller ( $10^{-2} \text{ MJ m}^{-2} \text{ s}^{-1}$ ) than that of the centimetre-sized rock specimens and without showing the reduction in the intermediate work rate. Such differences are not likely to be explained by the differences in experimental conditions (see Supplementary Information section 3).

Figure 3a and b shows the macroscopic friction data of experiments LB05-004 and LB04-007, respectively. The work rate of LB05-004 is

<sup>1</sup>National Research Institute for Earth Science and Disaster Prevention (NIED), Tsukuba 305-0006, Japan. <sup>2</sup>Central Research Institute of Electric Power Industry (CRIEPI), Abiko 270-1194, Japan.

<sup>3</sup>College of Science and Engineering, Ritsumeikan University, Kusatsu 525-8577, Japan.





**Figure 1 | Experimental system and results.** **a**, Schematic diagram of the large-scale biaxial experimental apparatus. The lower rock specimen (2.0 m long) was fixed on the large-scale shaking table and moved with it. The upper rock specimen (1.5 m long) was laid on the lower specimen and was held in place by a reaction force bar, which was fixed to the stationary outer reaction force floor. A normal load was applied by hydraulic jacks and measured with load cells serially connected with the jacks. The shear load was measured with a load cell fixed at the edge of the reaction force bar. The macroscopic friction coefficient is obtained from the measured shear load divided by the normal load. Local shear stresses were estimated from the shear strain measured with a strain gauge array (white squares),

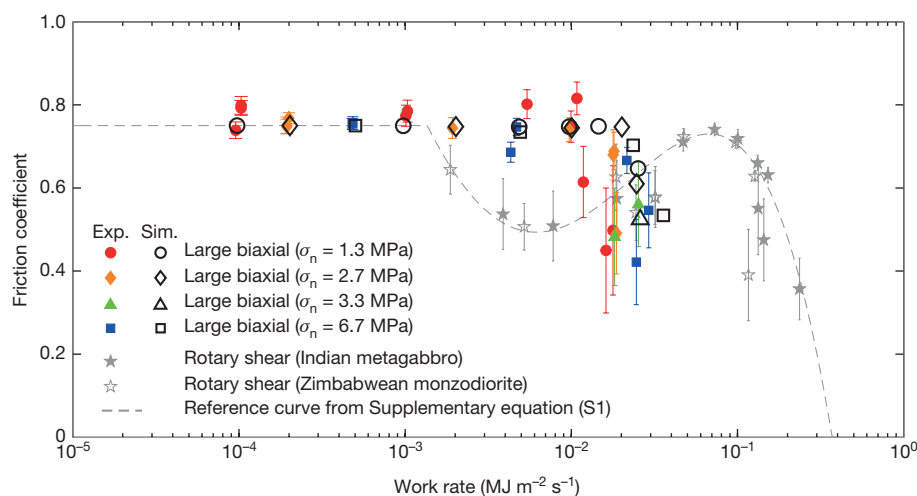
which was glued on the side face of the lower specimen along the fault. **b**, Photograph of a simulated fault surface just after an experiment. Localized damage (that is, grooves) and gouge materials (which appear as light grey lines) were generated by the frictional slip. Areas enclosed by yellow rectangles are shown in **c** and **d**. **c**, Close-up photograph of the beginning edge of a groove. The edge is very sharp. **d**, Close-up photograph of the ending edge of a groove. The groove was filled with the gouge materials, which looked heavily comminuted and partly consolidated. The relative slip direction of the upper specimen is denoted by the white arrows in **c** and **d**.

larger than that of LB04-007 since the loading rate of LB05-004 is higher (Supplementary Table 2). Figure 3c–f shows the evolution of the spatial distribution and its standard deviation of local shear stresses, which were measured by a local shear strain gauge array (see Fig. 1a and section 1 in Supplementary Information). We observed local shear stress concentrations and a related increase in spatial heterogeneity (standard deviation of local shear stress) with the fault slip during LB05-004, as illustrated in Fig. 3c and e. In contrast, we did not observe such an increase in local stress variation during LB04-007, in which the friction coefficient did not drop with slip (Fig. 3b, d, f). These observations suggest that the decrease in the macroscopic friction of metre-sized rocks is associated with the increase in spatial stress heterogeneity as well as the work rate. Similar results can be seen in other experiments (Supplementary Fig. 7).

We also evaluated the degree of spatial heterogeneity on the fault by mapping newly generated grooves and gouge using the photo image

data. A negative correlation was observed between the final friction coefficient and the spatial heterogeneity; the friction coefficients tended to be low when the grooves and gouge were heterogeneously generated during the experiments (see Supplementary Information section 5 and Supplementary Fig. 8). We found clear evidence of melting in the gouge collected in the large-scale biaxial experiments (see Supplementary Information section 5 and Supplementary Fig. 9), though frictional melting did not occur in centimetre-scale rotary shear experiments under similar mechanical conditions<sup>12</sup>. This material observation suggests intense stress concentrations with locally high work rate.

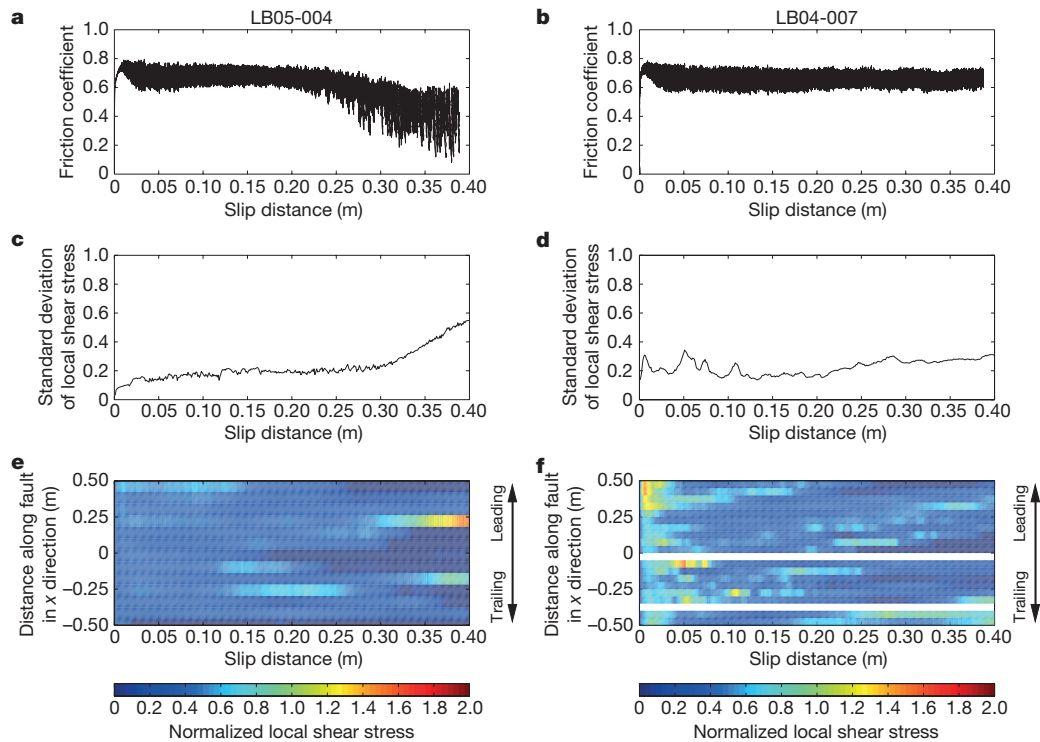
From these observations, we propose a stress localization model in which the contacting fault surfaces are composed of patch and off-patch areas with rather high and low normal stresses, respectively. The high normal stress on a patch causes high shear stress, high mechanical work and a high work rate upon loading, which lead to a high production rate of wear material (that is, gouge). The gouge generated is dragged by



**Figure 2 | The friction coefficient as a function of the work rate.**

The grey star symbols show the results of rotary shear experiments with centimetre-sized metagabbro and monzodiorite specimens, which were collected in India and Zimbabwe, respectively. The dashed line is a reference friction determined from those data and was used for the numerical simulation. Coloured symbols show the results of the large-scale biaxial experiments with metre-sized Indian metagabbro specimens. The friction starts to decrease at a work rate one order of magnitude smaller than that of the centimetre-sized rock specimens, without the reduction in

the intermediate work rate. 28 large-scale biaxial and 21 rotary shear experiments data are plotted, and are listed in Supplementary Tables 2 and 3, respectively. The centre value and the length of the error bars in those data represent the average and the one standard deviation of the friction coefficient, respectively, which are obtained during a certain slip interval for each experiment (see Supplementary Information section 1). Black open symbols represent the simulated macroscopic friction. The work-rate dependence of the simulated friction was fairly consistent with that of the biaxial experiments.



**Figure 3 | Slip-dependent evolution of friction and local shear stress heterogeneity during experiments.** **a, b**, Friction coefficients. **c, d**, Total amounts of normalized deviations of each local shear stress from the average. **e, f**, Local shear stresses normalized by input normal stress during experiments LB05-004 and LB04-007, respectively. The local shear stresses in the central part of the fault are shown here. Distances in

the  $x$  direction are measured from the centre of the lower specimen. The friction decreased with the increase in the local shear stress deviation during LB05-004, whereas both of them were almost constant during LB04-007. The 'leading' direction refers to the direction from which loading is imposed by the shaking table, and 'trailing' direction means the side where the reaction force bar is located (Fig. 1a).

the dislocating fault surface, which will cause further extension of the grooves, and thus will produce further gouge and stress heterogeneities. This is because the gouge layer on the groove (patch) will become thicker than the gouge layer on the surrounding area (off-patch). Therefore, these processes form a positive-feedback loop.

As a result of high work rate in the patch areas, the local frictional strength decreases rapidly with subsequent slip. Since most of the macroscopic shear load is sustained by the patch areas, the macroscopic frictional strength will be dominated by the rapid reduction of frictional strength in the patch areas. We note that the decrease in the frictional strength decelerates the gouge production, which causes a negative feedback. Therefore, these two opposite feedback effects compete during fault slips and so the macroscopic frictional behaviour depends on the competition. We also note that the positive-feedback effect enhances spatial stress heterogeneity as the fault slips even if the initial heterogeneity is small before the slip. In fact, deviation of the stress on the fault before the experiment was small in the present study (see section 3 of the Supplementary Information).

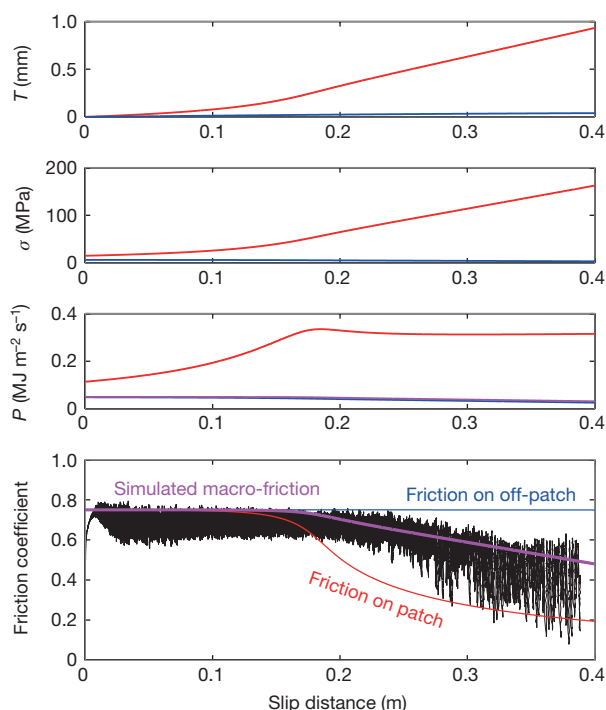
To quantitatively verify this idea, we numerically simulated a slip-evolved friction for both the patch and the off-patch areas, as well as the macroscopic friction. Calculations were done according to the flow chart shown in Supplementary Fig. 10. In this computation, we assumed that the local friction coefficients on both the patch and off-patch areas follow the work-rate-dependent friction law estimated from the centimetre-scale rotary shear experiments (Fig. 2). We also assume that the local friction does not immediately attain the target friction coefficient determined by the work rate, but gradually approaches it until the breakdown work is equal to the fracture energy ( $E_f$ ), which is estimated from the centimetre-scale rotary shear experiments. Therefore, the friction on the patch and off-patch areas ( $\mu_*$ ) is written as  $\mu_* = F_1(\mu_*, \sigma_*, V, s, E_f)$ , where  $\sigma_*$  is the local normal stress,  $V$  is the slip rate, and  $s$  is the slip distance. The asterisk denotes either 'p' or 'op', for the patch or off-patch area, respectively. The expression for  $F_1$  can

be found in section 6 of the Supplementary Information (as for the below-mentioned  $F_2$ ,  $F_3$  and  $F_4$ ).

The normal stresses on the patch and off-patch areas ( $\sigma_*$ ) depend on the thickness of the gouge layer as well as the area ratio:  $\sigma_* = F_2(\sigma, R_{Ap}, \Delta T)$ , where  $\sigma$  is the macroscopic normal stress,  $R_{Ap}$  is the fraction of the patch areas over the entire fault area, and  $\Delta T$  is the difference between the thicknesses of the gouge layer in the patch and off-patch areas. We estimated  $R_{Ap}$  using the areas occupied with the gouge (which were obtained from the image analyses), and assumed that  $R_{Ap}$  was constant during the experiments. The thickness of the gouge layer ( $T_*$ ) is a function of the input frictional work and the wear rate ( $R_{w*}$ ) as follows:  $T_* = F_3(\mu_*, \sigma_*, s, R_{w*})$ .  $R_{w*}$  was evaluated from the mass of the collected gouge materials per unit input work. The variables were updated at each step and the macroscopic friction ( $\mu_{macro}$ ) was calculated from the equation  $\mu_{macro} = F_4(\mu_p, \mu_{op}, \sigma_p, \sigma_{op}, \sigma, R_{Ap})$ .

The detailed procedures of the simulation, including how the related parameters were determined, are described in section 6 of the Supplementary Information. The simulations successfully reproduced a concentration of normal stress and work rate on the patch, and also a weakening in the macroscopic friction with the slip that fits well with the selected experimental data (Fig. 4). By conducting simulations under various conditions, which are the same as those for the present biaxial experiments, we confirmed that the work-rate dependence of the simulated friction was fairly consistent with that of the biaxial experiments (Fig. 2). From these results, we also confirmed that the slip-evolved spatial heterogeneity caused the scale-dependent rock friction.

Rock samples collected from fault zones are frequently investigated in the laboratory in an attempt to understand the frictional behaviour of natural faults<sup>21,22</sup>. However, the present study suggests that the estimated frictional properties might not properly represent the behaviour of larger, natural faults because spatial heterogeneity, which is common or even dominant in nature, can greatly affect the macroscopic friction.



**Figure 4 | Comparison of experimental data and simulation results.** Slip-dependent evolution of friction coefficients and calculated parameters.  $T$  is the gouge layer thickness,  $\sigma$  is the normal stress, and  $P$  is the work rate. The black line in the bottom panel shows the experimental data of LB05-004. The purple line shows the simulated macroscopic parameter. Red and blue lines indicate the simulated parameters on the patch areas and the off-patch areas, respectively.

The amount of gouge production on the patch area was a key process in the present study, but frictional melting may be dominant under natural conditions. Even in such a melt-dominant situation, a similar mechanism to that presented here will work, as shown by an earlier investigation<sup>13</sup>.

We also note that the normal stress concentration can cause not only the decrease in dynamic friction but also an unstable fault slip by reducing the critical nucleation length  $L_c$ . If we assume a friction law in which frictional force is roughly proportional to normal force—such as the rate- and state-dependent friction law—then  $L_c$  will be inversely proportional to the normal stress<sup>23,24</sup>. If the area of stress concentration is large enough relative to the diminished  $L_c$  and the frictional properties are uniform there, unstable slip will start to occur in that area. Small foreshocks during slow slip before a main shock have been observed in laboratory experiments and on natural faults<sup>18,25–28</sup>. These foreshocks should be related to the spatial heterogeneity, as observed in the present study. We conclude that spatial stress heterogeneity can facilitate the occurrence of an earthquake.

In the present large-scale experiments, we found that the rock friction is scale-dependent, and we believe that spatial stress heterogeneity and the resulting localized high work rate cause this dependence. Although some earlier work<sup>13,29,30</sup> qualitatively suggested the idea that local frictional phenomena can control the macroscopic friction, here we have quantitatively confirmed it using metre-scale experiments. If large rocks lose their frictional strength at a small work rate, a natural fault may rapidly weaken at an unexpectedly early stage during slip. More studies on large-scale friction experiments, increasing normal stress for example, are needed to explore scale dependence further and to understand better the frictional properties of natural faults.

Received 23 April; accepted 6 October 2015.

1. Tsutsumi, A. & Shimamoto, T. High-velocity frictional properties of gabbro. *Geophys. Res. Lett.* **24**, 699–702 (1997).

2. Goldsby, D. L. & Tullis, T. E. Low frictional strength of quartz rocks at subseismic slip rates. *Geophys. Res. Lett.* **29**, 1844, <http://dx.doi.org/10.1029/2002GL015240> (2002).
3. Goldsby, D. L. & Tullis, T. E. Flash heating leads to low frictional strength of crustal rocks at earthquake slip rates. *Science* **334**, 216–218 (2011).
4. Di Toro, G., Goldsby, D. L. & Tullis, T. E. Friction falls towards zero in quartz rock as slip velocity approaches seismic rates. *Nature* **427**, 436–439 (2004).
5. Di Toro, G. *et al.* Fault lubrication during earthquakes. *Nature* **471**, 494–498 (2011).
6. Hirose, T. & Shimamoto, T. Growth of molten zone as a mechanism of slip weakening of simulated faults in gabbro during frictional melting. *J. Geophys. Res.* **110**, B05202 (2005).
7. Han, R., Shimamoto, T., Hirose, T., Ree, J. H. & Ando, J. Ultralow friction of carbonate faults caused by thermal decomposition. *Science* **316**, 878–881 (2007).
8. Han, R., Hirose, T. & Shimamoto, T. Strong velocity weakening and powder lubrication of simulated carbonate faults at seismic slip rates. *J. Geophys. Res.* **115**, B03412 (2010).
9. Han, R., Hirose, T., Shimamoto, T., Lee, Y. & Ando, J. Granular nanoparticles lubricate faults during seismic slip. *Geology* **39**, 599–602 (2011).
10. Nielsen, S., Di Toro, G., Hirose, T. & Shimamoto, T. Frictional melt and seismic slip. *J. Geophys. Res.* **113**, B01308 (2008).
11. Reches, Z. & Lockner, D. A. Fault weakening and earthquake instability by powder lubrication. *Nature* **467**, 452–455 (2010).
12. Mizoguchi, K. & Fukuyama, E. Laboratory measurements of rock friction at subseismic slip velocities. *Int. J. Rock Mech. Min. Sci.* **47**, 1363–1371 (2010).
13. Brown, K. M. & Fialko, Y. ‘Melt welt’ mechanism of extreme weakening of gabbro at seismic slip rates. *Nature* **488**, 638–641 (2012).
14. Ohnaka, M. & Shen, L. F. Scaling of the shear rupture process from nucleation to dynamic propagation: implications of geometric irregularity of the rupturing surfaces. *J. Geophys. Res.* **104**, 817–844 (1999).
15. Dieterich, J. H. Preseismic fault slip and earthquake prediction. *J. Geophys. Res.* **83**, 3940–3948 (1978).
16. Yoshida, S. & Kato, A. Single and double asperity failures in a large-scale biaxial experiment. *Geophys. Res. Lett.* **28**, 451–454 (2001).
17. Beeler, N. M. *et al.* Observed source parameters for dynamic rupture with nonuniform initial stress and relatively high fracture energy. *J. Struct. Geol.* **38**, 77–89 (2012).
18. McLaskey, G. C. & Kilgore, B. D. Foreshocks during the nucleation of stick-slip instability. *J. Geophys. Res.* **118**, 2982–2997 (2013).
19. Fukuyama, E. *et al.* Large-scale Biaxial Friction Experiments using a NIED Large-Scale Shaking Table—Design of Apparatus and Preliminary Results. Report 81 of the National Research Institute for Earth Science and Disaster Prevention, 15–35, [http://dil-opac.bosai.go.jp/publication/nied\\_report/PDF/81/81-3fukuyama.pdf](http://dil-opac.bosai.go.jp/publication/nied_report/PDF/81/81-3fukuyama.pdf) (2014).
20. Byerlee, J. D. Friction of rocks. *Pure Appl. Geophys.* **116**, 615–626 (1978).
21. Lockner, D. A., Morrow, C., Moore, D. & Hickman, S. Low strength of deep San Andreas fault gouge from SAFOD core. *Nature* **472**, 82–85 (2011).
22. Ujii, K. *et al.* Low coseismic shear stress on the Tohoku-Oki megathrust determined from laboratory experiments. *Science* **342**, 1211–1214 (2013).
23. Dieterich, J. H. In *Earthquake Source Mechanics* (eds Das, S., Boatwright, J. & Scholz, C. H.) 37–47 (AGU, 1986).
24. Dieterich, J. H. Earthquake nucleation on faults with rate- and state-dependent strength. *Tectonophysics* **211**, 115–134 (1992).
25. McLaskey, G. C. & Lockner, D. A. Preslip and cascade processes initiating laboratory stick slip. *J. Geophys. Res.* **119**, 6323–6336 (2014).
26. Bouchon, M. *et al.* Extended nucleation of the 1999  $M_w$  7.6 Izmit earthquake. *Science* **331**, 877–880 (2011).
27. Bouchon, M., Durand, V., Marsan, D., Karabulut, H. & Schmittbuhl, J. The long precursory phase of most large interplate earthquakes. *Nature Geosci.* **6**, 299–302 (2013).
28. Kato, A. *et al.* Propagation of slow slip leading up to the 2011  $M_w$  9.0 Tohoku-Oki earthquake. *Science* **335**, 705–708 (2012).
29. Boneh, Y., Chang, J. C., Lockner, D. A. & Reches, Z. Evolution of wear and friction along experimental faults. *Pure Appl. Geophys.* **171**, 3125–3141 (2014).
30. Boneh, Y., Sagy, A. & Reches, Z. Frictional strength and wear-rate of carbonate faults during high-velocity, steady-state sliding. *Earth Planet. Sci. Lett.* **381**, 127–137 (2013).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank A. Schubnel and F. X. Passelègue for help with some of the experiments. We thank G. C. McLaskey for comments. Reviews by Z. Reches improved the manuscript. This study was supported by the NIED research project entitled “Development of monitoring and forecasting technology for crustal activity” and JSPS KAKENHI grant number 23340131.

**Author Contributions** E.F., F.Y., K.M. and H.K. designed and developed the large-scale friction experiments. F.Y. performed the analysis of the experimental data and numerical modelling. S.T. analysed the collected gouge material. F.Y. wrote the manuscript, with contributions from E.F. and S.X. All authors contributed to the performance of the large-scale friction experiments.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to F.Y. ([yamafuto@bosai.go.jp](mailto:yamafuto@bosai.go.jp)).



# The ontogeny of fairness in seven societies

P. R. Blake<sup>1\*</sup>, K. McAuliffe<sup>2,3,4\*</sup>, J. Corbit<sup>5</sup>, T. C. Callaghan<sup>6</sup>, O. Barry<sup>7</sup>, A. Bowie<sup>4,8</sup>, L. Kleutsch<sup>8</sup>, K. L. Kramer<sup>9</sup>, E. Ross<sup>4</sup>, H. Vongsachang<sup>4,8</sup>, R. Wrangham<sup>4</sup> & F. Warneken<sup>8</sup>

**A sense of fairness plays a critical role in supporting human cooperation<sup>1–3</sup>. Adult norms of fair resource sharing vary widely across societies, suggesting that culture shapes the acquisition of fairness behaviour during childhood<sup>4,5</sup>. Here we examine how fairness behaviour develops in children from seven diverse societies, testing children from 4 to 15 years of age ( $n = 866$  pairs) in a standardized resource decision task<sup>6,7</sup>. We measured two key aspects of fairness decisions: disadvantageous inequity aversion (peer receives more than self) and advantageous inequity aversion (self receives more than a peer). We show that disadvantageous inequity aversion emerged across all populations by middle childhood. By contrast, advantageous inequity aversion was more variable, emerging in three populations and only later in development. We discuss these findings in relation to questions about the universality and cultural specificity of human fairness.**

Fairness requires a concern for relative payoffs between oneself and others or between third parties. A key component of the human sense of fairness is inequity aversion, defined as the willingness to sacrifice material payoffs for the sake of greater equality<sup>2,8</sup>. Importantly, inequity aversion takes two forms, depending on which side of inequality one faces. Disadvantageous inequity aversion (DI) occurs to avoid receiving less than a peer, and advantageous inequity aversion (AI) occurs to avoid receiving more than a peer. Both DI and AI have been suggested to support cooperation within societies<sup>2,8</sup>. DI entails an immediate cost but can provide long-term benefits by preventing competitors from attaining a relative advantage and signalling that one will not tolerate being exploited<sup>8</sup>. AI entails a larger immediate sacrifice by rejecting a relative advantage. It may signal that one is a good cooperative partner who will not exploit others. Given the larger immediate cost, AI expresses a stronger sense of fairness.

The distinction between DI and AI has emerged as a milestone for research on fairness across many disciplines. Behavioural economics experiments of various types show that although adults exhibit both forms of inequity aversion, they are more willing to pay a cost to avoid a disadvantage (DI) than an advantage (AI)<sup>1,2,9</sup>. Neuroscientific studies indicate that DI and AI are supported by different neural processes<sup>10</sup>. Animal studies suggest that DI may exist in nonhuman primates and other social species, whereas AI is more rare, possibly occurring only in humans<sup>8</sup>. This suggests that the psychology behind DI may have deep evolutionary roots, while AI may depend on factors found only in our species.

Research with children in the West has revealed different developmental patterns for DI and AI and different motivations for each behaviour<sup>6,11–15</sup>. By 4 years of age, children show DI, which appears to be motivated by spite<sup>7</sup>. By contrast, AI appears at about 8 years of age. This indicates that a strong egalitarian preference is characteristic only of later development, potentially because children have internalized social norms that guide their behaviour<sup>16,17</sup>. However, it is unknown

whether the separate developmental trajectories of DI and AI are consistent across different populations, and perhaps even universal aspects of human behaviour. Testing this requires cross-cultural comparisons, yet most research on the development of fairness has been conducted with children from societies that are Western, educated, industrialized, rich and democratic (WEIRD)<sup>18</sup>. This is problematic because it ignores potentially important cross-cultural variation in the development of fairness. Further, studies using a range of economic tasks with adults have found that WEIRD societies tend to be outliers on many measures of fairness and cooperation<sup>4,5,19,20</sup>.

To date, no study, to our knowledge, has examined the development of inequity aversion across societies. Recent research on the development of prosociality in different societies has found a general increase in generosity with age<sup>17,21</sup>. Other studies have found that children in some non-Western societies tend to share resources more equally compared to children in the West<sup>22,23</sup>. Experiments with children in two non-Western populations found that participants recognize adult norms for AI<sup>24,25</sup>. However, no study has investigated how both forms of inequity aversion emerge during development across different societies.

Here we describe the first developmental study of inequity aversion across societies. We used the inequity game, an intuitive experimental task that measures both DI and AI, differentiates between generosity and fairness as motives, and has been validated for Western children from 4 to 9 years of age<sup>6,26</sup>. In the inequity game, two children sat across from each other at an apparatus, randomly assigned to either disadvantageous (DI condition) or advantageous (AI condition) allocations (see Fig. 1). We used small food treats as rewards. One child, the actor, had a choice between accepting the allocation or rejecting it. The recipient played no part in the decision.

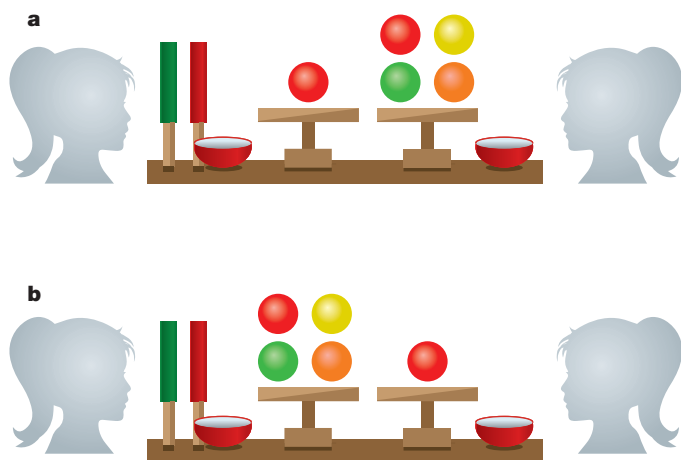
An experimenter presented a sequence of equal (1–1) and unequal (1–4, DI; 4–1, AI) allocations of food rewards to the pair of children. Rejections resulted in a 0–0 payoff and thus went against immediate self-interest and were not altruistic or generous. Rejecting unequal allocations more often than the equal allocations provided a measure of inequity aversion. Children were recruited in seven countries representing a diverse set of communities (see Table 1). Two communities fit the WEIRD characterization and five were non-WEIRD. We tested a total of  $n = 866$  actors from 4 to 15 years old, each paired with a same-gender peer of a similar age.

Our main goal was to measure for the first time the presence and the developmental emergence of DI and AI across a range of societies. We hypothesized that DI would be more common across populations than AI. Given that adults from WEIRD societies have a stronger tendency, relative to non-WEIRD societies, to adhere to norms of equality even when it is costly to the self<sup>4,5,19</sup>, we hypothesized that children from the WEIRD populations in our sample would be more likely to show AI than children from the other societies. As existing studies with Western

<sup>1</sup>Department of Psychological & Brain Sciences, Boston University, Boston, Massachusetts 02215, USA. <sup>2</sup>Department of Psychology, Yale University, New Haven, Connecticut 06520, USA.

<sup>3</sup>Department of Psychology, Boston College, Chestnut Hill, Massachusetts 02467, USA. <sup>4</sup>Department of Human Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA. <sup>5</sup>Department of Psychology, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada. <sup>6</sup>Faculty of Arts and Sciences, Department of Psychology, St Francis Xavier University, Antigonish, Nova Scotia B2G 2W5, Canada. <sup>7</sup>University of Cheikh Anta Diop, Faculty of Science and Technology for Education and Training, BP 5036 Dakar Fann, Senegal. <sup>8</sup>Department of Psychology, Harvard University, Cambridge, Massachusetts 02138, USA. <sup>9</sup>Department of Anthropology, University of Utah, Salt Lake City, Utah 84112, USA.

\*These authors contributed equally to this work.



**Figure 1 | Experimental set-up.** **a, b,** Illustration of the apparatus and allocations for disadvantageous inequity (**a**) and advantageous inequity (**b**). The actor is on the left.

children have consistently found different developmental trajectories for DI and AI, we predicted that where both behaviours were found, DI would emerge earlier than AI.

In line with our first two hypotheses, our overall results indicated that DI was present in all populations tested, whereas AI was present in only three of the seven populations. AI was found in our two WEIRD societies (USA and Canada), but also in Uganda, a result we discuss below. As predicted, DI and AI showed different developmental trajectories, with DI emerging early in childhood and AI later, if at all.

We now turn to a more detailed examination of these results. In all seven populations, children sacrificed a food reward to prevent a peer from receiving more than them. Thus DI emerged during childhood in societies ranging from small villages with a subsistence economy to large industrialized cities (Fig. 2). This result suggests that DI is a more general feature of human behaviour. However, there was also variability in the age of emergence and the strength of the effect, indicating that cultural factors may influence the expression of DI. Children rejected a disadvantage earliest in the USA and Canada (by 4–6 years of age) and latest in Mexico (by 10 years of age). In all groups except Mexico, DI became stronger with age.

The consistent appearance of DI in all populations is notable because this behaviour denies rewards both to the self and the peer recipient. Rejections of DI are thus both costly and not prosocial. There are two plausible interpretations of this result. One possibility is that DI represents an early sense of fairness. Preschool children in the USA are aware of norms of equal distribution<sup>16</sup>, but they are reluctant to place themselves at a disadvantage even when they are less deserving than a partner and become visibly upset when they receive less than a peer<sup>26,27</sup>. DI may thus represent an early application of norms of fairness with a focus on unfairness to oneself. However, under this interpretation children in Mexico would be less concerned with norms of fairness or have a very different understanding of what fairness means.

Alternatively, DI may have little to do with norms of fairness and more to do with preserving one's status relative to potential competitors<sup>7,8</sup>. Under this hypothesis, one would expect cultural norms of competition to enhance DI during childhood. Specifically, in societies with stronger norms of childhood competition, DI should appear at younger ages compared to societies with weaker childhood competition. One factor that could affect competition among children is the specific relationship between the pairs in our sample. In the majority of the populations in our sample, the nature of the relationship between pairs of children varied between children who knew each other (familiar) and those who had not met before (unfamiliar) (Supplementary Table 2). The high proportion of familiar pairs tested within a relatively small population in Mexico may reflect a more close-knit community for the Maya in which children experience less competition and envy or more generosity towards peers.

While DI was common, AI was found in only three of our communities: USA, Canada and Uganda (Fig. 3). In these societies, rejections of advantageous allocations increased with age with AI appearing to emerge by pre-adolescence (Supplementary Information). AI has been viewed as an expression of a norm-based sense of fairness with a focus on unfairness to others<sup>26</sup>. Given that Western societies tend to emphasize establishing and enforcing norms of equality<sup>19,20</sup>, it is possible that children in these communities face social pressures to internalize and enact these norms earlier in development compared to other societies. Although Uganda is a non-Western society, the schools from which we recruited children interacted frequently with Western teachers and researchers whose contributions could in theory have led to the promotion of Western norms of fairness. However, while this explanation is plausible, it remains possible that children in Uganda reject an advantage for other reasons not linked to Western norms. If this is the case, we would expect to see AI emerge in children in other communities in Uganda with similar cultural norms but different institutional structures.

Our final prediction concerned the timing of development of DI and AI in relation to each other. In the three societies in which both forms of inequity aversion appeared, DI emerged several years earlier than AI. The relatively early emergence of DI and the substantial delay in the development of AI is in line with the notion that these two forms of inequity aversion are supported by different psychological processes. The primacy of DI in development is also consistent with the view that a self-focused form of fairness precedes the other-focused form of AI. This progression may represent a shift from a concern for one's own prospects to a concern for the greater good, or at least to a concern for the consequences of appearing unfair.

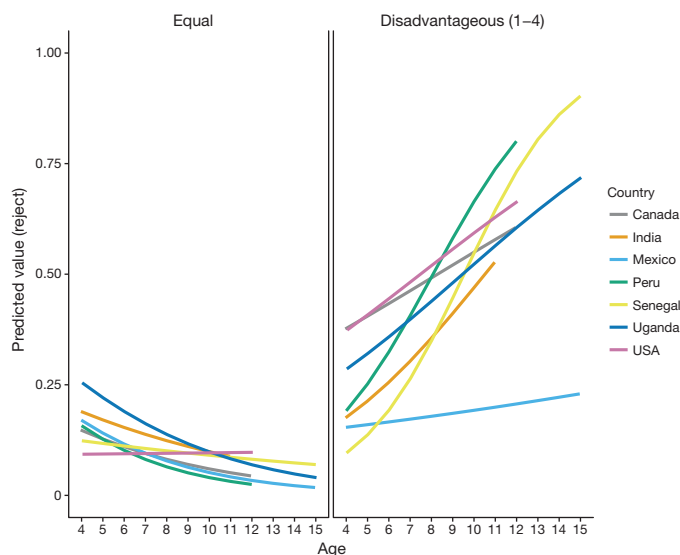
The comparison of the development of DI and AI across communities offers insights into the potential role of culture in the ontogeny of fairness. First, DI appeared during childhood in all seven of the societies tested, but the developmental trajectory varied across the sites. One interpretation of this result is that while children have a general tendency to develop DI, cultural factors can shape its development. This consistent tendency probably reflects the fact that DI is beneficial in the long run, regardless of local norms. In contrast to DI, AI appeared during childhood in only three of the societies

**Table 1 | Descriptive summary of testing sites**

Site	Locale	Pop. size*	Dominant language	Dominant religion	Economy
Canada	Antigonish	4k	English	Catholic (56%)	Professional, trade/service, agriculture
India	Villages, Andhra Pradesh	2k	Telegu	Hindu (63%)	Agriculture, labour
Mexico	Xculoc, Puuc region	500	Maya	Catholic (90%)	Agriculture, labour
Peru	Villages, San Pedro de Saño	700	Spanish	Catholic (93%)	Agriculture, labour
Senegal	Dakar	3m	Wolof	Muslim (94%)	Trade/service, labour, fishing
Uganda	Villages, Fort Portal	500	Rutooro	Catholic and Anglican (78%)	Agriculture, labour
USA	Boston	646k	English	Protestant and Catholic (75%)	Professional, trade/service, labour

Populations shown are the average size for villages tested in India, Peru and Mexico.

k = thousand, m = million.



**Figure 2 | Disadvantageous inequity aversion.** Estimates of children's rejections of equal (left panel) and disadvantageous (right panel) allocations across ages and countries from a generalized linear mixed model (GLMM) controlling for gender and trial effects. Lines are truncated based on the age range of children tested in each country. Total actors,  $n = 429$ .

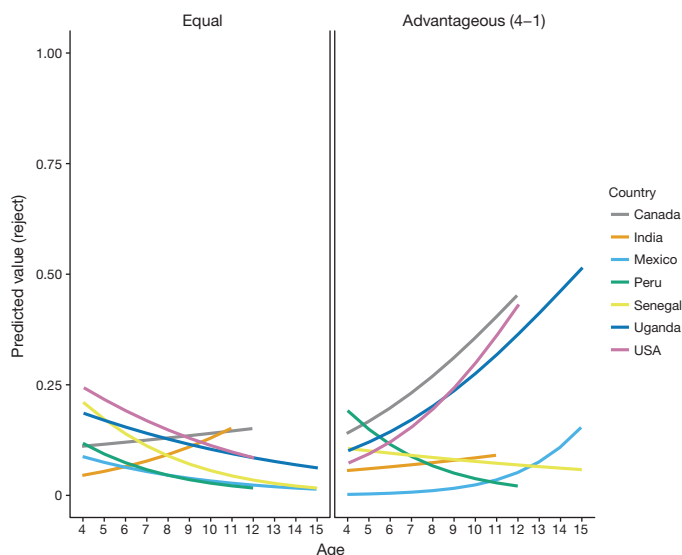
tested. This suggests that culture has a more potent influence on this form of fairness. Cultural input may be required to encourage the emergence of AI in childhood because this behaviour goes strongly against immediate self-interest.

Second, it is notable that the AI emerged during late childhood in the USA and Canada, our two WEIRD populations. Although further tests are needed to identify the social factors that make these sites conducive to the emergence of AI, we highlight two factors that could contribute to this pattern. Parents from Western societies are known to more strongly encourage children to show autonomy and independence, which may result in an earlier concern for one's own social standing and reputation in peer groups<sup>28</sup>. Children in these populations may thus engage in AI with peers in order to maintain a reputation as a good cooperator. If this is true, children may be rejecting an advantage in order to serve their longer-term self-interest.

An alternative explanation is that a community's degree of economic exchange among unfamiliar individuals influences the developmental emergence of AI. So-called market integration has been found to predict adult fairness behaviours across a range of diverse societies<sup>4,20</sup>. Specifically, societies with more pervasive market interactions tend to have stronger norms for equal distribution of resources as measured by donations to others in the dictator game and offers proposed in the ultimatum game<sup>5,20</sup>. This effect is particularly robust in Western societies, suggesting that children in these populations are more often exposed to consistent egalitarian norms. If this is the case, children in these environments may be more likely to internalize a norm of AI once they are old enough to adopt adult behaviours.

One potential concern is that differences in the reward value used at each site could have influenced children's decisions. However, additional analyses show that this does not explain our results, as the pattern of children's responses was the same when we compared their decisions for higher and lower value rewards within societies (see Supplementary Information). Moreover, the same rewards were used for tests of DI and AI, and thus any difference in rejections between conditions cannot be explained by the type of resource alone.

The current study highlights several important avenues for future research. First, although our study focused on the developmental emergence of the two forms of inequity aversion, these behaviours may change further across the lifespan. Indeed, evidence from Western



**Figure 3 | Advantageous inequity aversion.** Estimates of children's rejections of equal (left panel) and advantageous (right panel) allocations across ages and countries from a GLMM controlling for gender and trial effects. Lines are truncated based on the age range of children tested in each country. Total actors,  $n = 437$ .

adults suggests that the strength of both DI and AI may vary over adulthood<sup>15,29,30</sup>. Thus, although AI was absent in childhood in the majority of populations tested, this does not imply that it is absent in adulthood. Children in many societies may not adopt this particular fairness behaviour until adolescence or adulthood.

Second, additional insights can be gained by testing an even greater diversity of populations. Although our sample included populations which varied in factors such as population size, economic structure and major religions, we did not test societies of foragers, horticulturalists or pastoralists. Studies with adults from these societies have found marked variation in fairness behaviour compared to other populations<sup>5</sup>. Understanding how the ontogeny of DI and AI vary across a more complete spectrum of human societies would offer a stronger test of the possibility that DI is a universal feature of human behaviour and help to identify the circumstances under which different forms of inequity aversion emerge.

Finally, future work should explore the specific factors that influence the development of inequity aversion both between and within populations. Isolating the characteristics of societies that can explain variation between populations remains an important goal and will require systematic ethnographic measures of values, norms and ecologies within the communities in addition to experiments. Within-population experiments would also help to identify how specific social influences, such as formal schooling or particular social relationships, affect inequity aversion during childhood. In addition, learning studies in which children are systematically exposed to different fairness behaviours could provide insight into the extent to which DI and AI are influenced by cultural transmission<sup>3</sup>. Future research should combine our experimental-developmental approach with cultural-level analyses to further explain the factors that give rise to the human sense of fairness.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 1 July; accepted 7 September 2015.**

**Published online 18 November 2015.**

1. Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R. & Smirnov, O. Egalitarian motives in humans. *Nature* **446**, 794–796 (2007).
2. Fehr, E. & Schmidt, K. M. A theory of fairness, competition, and cooperation. *Q. J. Econ.* **114**, 817–868 (1999).



3. Chudek, M. & Henrich, J. Culture-gene coevolution, norm-psychology and the emergence of human prosociality. *Trends Cogn. Sci.* **15**, 218–226 (2011).
4. Henrich, J. *et al.* “Economic man” in cross-cultural perspective: behavioral experiments in 15 small-scale societies. *Behav. Brain Sci.* **28**, 795–815 (2005).
5. Henrich, J. *et al.* Costly punishment across human societies. *Science* **312**, 1767–1770 (2006).
6. Blake, P. R. & McAuliffe, K. “I had so much it didn’t seem fair”: Eight-year-olds reject two forms of inequity. *Cognition* **120**, 215–224 (2011).
7. McAuliffe, K., Blake, P. R. & Warneken, F. Children reject inequity out of spite. *Biol. Lett.* **10**, 20140743 (2014).
8. Brosnan, S. F. & de Waal, F. B. Evolution of responses to (un)fairness. *Science* **346**, 1251776 (2014).
9. Loewenstein, G. F., Thompson, L. & Bazerman, M. Social utility and decision making in interpersonal contexts. *J. Pers. Soc. Psychol.* **57**, 426–441 (1989).
10. Güroğlu, B., Will, G. J. & Crone, E. A. Neural correlates of advantageous and disadvantageous inequity in sharing decisions. *PLoS ONE* **9**, e107996 (2014).
11. McAuliffe, K., Blake, P. R., Kim, G., Wrangham, R. W. & Warneken, F. Social influences on inequity aversion in children. *PLoS ONE* **8**, e80966 (2013).
12. Sheskin, M., Bloom, P. & Wynn, K. Anti-equality: social comparison in young children. *Cognition* **130**, 152–156 (2014).
13. Leimgruber, K. L., Shaw, A., Santos, L. R. & Olson, K. R. Young children are more generous when others are aware of their actions. *PLoS ONE* **7**, e48292 (2012).
14. Fehr, E., Bernhard, H. & Rockenbach, B. Egalitarianism in young children. *Nature* **454**, 1079–1083 (2008).
15. Fehr, E., Glätzle-Rützler, D. & Sutter, M. The development of egalitarianism, altruism, spite and parochialism in childhood and adolescence. *Eur. Econ. Rev.* **64**, 369–383 (2013).
16. Smith, C. E., Blake, P. R. & Harris, P. L. I should but I won’t: why young children endorse norms of fair sharing but do not follow them. *PLoS ONE* **8**, e59510 (2013).
17. House, B. R. *et al.* Ontogeny of prosocial behavior across diverse societies. *Proc. Natl Acad. Sci. USA* **110**, 14586–14591 (2013).
18. Henrich, J., Heine, S. J. & Norenzayan, A. The weirdest people in the world? *Behav. Brain Sci.* **33**, 61–83, discussion 83–135 (2010).
19. Herrmann, B., Thoni, C. & Gächter, S. Antisocial punishment across societies. *Science* **319**, 1362–1367 (2008).
20. Henrich, J. *et al.* Markets, religion, community size, and the evolution of fairness and punishment. *Science* **327**, 1480–1484 (2010).
21. Rochat, P. *et al.* Fairness in distributive justice by 3- and 5-year-olds across seven cultures. *J. Cross Cult. Psychol.* **40**, 416–442 (2009).
22. Robbins, E. & Rochat, P. Emerging signs of strong reciprocity in human ontogeny. *Front. Psychol.* **2**, 353 (2011).
23. Schäfer, M., Haun, D. B. & Tomasello, M. Fair is not fair everywhere. *Psychol. Sci.* **26**, 416–442 (2015).
24. Paulus, M. Children’s inequity aversion depends on culture: a cross-cultural comparison. *J. Exp. Child Psychol.* **132**, 240–246 (2015).
25. Shaw, A. & Olson, K. R. Children discard a resource to avoid inequity. *J. Exp. Psychol. Gen.* **141**, 382–395 (2012).
26. Blake, P. R., McAuliffe, K. & Warneken, F. The developmental origins of fairness: the knowledge-behavior gap. *Trends Cogn. Sci.* **18**, 559–561 (2014).
27. LoBue, V., Nishida, T., Chiong, C., DeLoache, J. S. & Haidt, J. When getting something good is bad: even three-year-olds react to inequality. *Soc. Dev.* **20**, 154–170 (2011).
28. Keller, H. *et al.* Cultural models, socialization goals, and parenting ethnotheories: a multicultural analysis. *J. Cross Cult. Psychol.* **37**, 155–172 (2006).
29. Sutter, M. Outcomes versus intentions: on the nature of fair behavior and its development with age. *J. Econ. Psychol.* **28**, 69–78 (2007).
30. Bahry, D. L. & Wilson, R. K. Confusion or fairness in the field? Rejections in the ultimatum game under the strategy method. *J. Econ. Behav. Organ.* **60**, 37–54 (2006).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We are grateful to the following people for their help with this research: participating children, families, communities and schools; research assistants who collected and coded data; Harvard IQSS; J. Greene; P. Harris and our funding sources (Harvard Academy Junior Faculty Development Grant, Radcliffe Institute for Advanced Study (F.W.); Herchel Smith Harvard Undergraduate Science Research Program (A.B.); Harvard College Research Program (H.V.); John Templeton Foundation (P.R.B.); Harvard Department of Human Evolutionary Biology (K.M.).

**Author Contributions** P.R.B., K.M., and F.W. planned the study and performed the data analyses. P.R.B., K.M., J.C., T.C.C., O.B., A.B., L.K., K.L.K., E.R., H.V., R.W. and F.W. contributed to the research and writing the paper.

**Author Information** The data and code for the statistical analyses are stored in Dryad Data package title: The ontogeny of fairness in seven societies; <http://dx.doi.org/10.5061/dryad.g3925>. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.R.B. (pblake@bu.edu) or K.M. (mcaulikg@bc.edu).

## METHODS

All study procedures and protocols were approved by Harvard University's Institutional Review Board, IRB F18470-108, F18470-118, and F18643-105. Additional approvals for the same protocols were obtained for research teams not associated with Harvard, including the Institutional Review Board of St. Francis Xavier University, Antigonish, Canada (IRB number 21630), the University of Utah (IRB number 00065740), the Cheik Anta Diop University in Senegal and the Uganda National Council for Science and Technology (IRB number SS 2761). The IRB approved procedures were followed for informed parental consent, child assent and the study protocol.

**Subjects.** We tested 866 actors from 4 to 15 years from 7 countries (Canada, India, Mexico, Peru, Senegal, Uganda and USA), each paired with a same-gender peer of a similar age. See Supplementary Information for details concerning recruitment, consent, sample characteristics and site descriptions.

**Design.** We used a  $2 \times 2$  design with inequity type (DI or AI) as a between-subjects variable and distribution (equal or unequal) as a within-subjects variable. Each child was tested in 16 trials, half with an equal and half with an unequal distribution (either disadvantageous or advantageous, depending on inequity type). In order to ensure proper counterbalancing and sample size at different ages, we created age-range groupings before testing (4–6, 7–9, 10–12, 13–15 years of age). Within each age group, children were randomly assigned to either the DI or AI condition.

**Experimental procedures.** We used the inequity game<sup>6</sup>. Two children sit across from each other at an apparatus. One child is assigned to the position of the actor who makes decisions and the other child is a passive recipient. An experimenter places allocations of treats on two trays that are designated for the actor and recipient.

The actor then decides whether to accept or reject the allocation by pulling different handles: pulling a green handle accepts the allocation by tilting the trays outward and delivering the treats to both participants; pulling a red handle rejects the allocation by tilting the trays inward and dropping the treats into a middle bowl so that neither participant receives them. The inequity apparatus provides an intuitive way for children to understand the consequences of their decisions and has been tested with children as young as 4 years of age. Our design differs from standard behavioural economics experiments with adults, which maintain anonymity of the actor and recipient with respect to each other. For young child participants, however, understanding the consequences of their sharing behaviour with absent peers in anonymous interactions may be too taxing and uncommon in everyday life. Accordingly we adopted a more age-sensitive method, in which both participants are present, that is typical in cross-cultural studies with children<sup>17</sup>. See Supplementary Information for a full protocol.

**Statistical methods.** Our dependent measure was children's decision to reject ( $=1$ ) or accept ( $=0$ ) a given allocation (see Supplementary Fig. 2). We used multilevel logistic regressions to assess whether children were more likely to reject unequal over equal distributions, and whether this pattern differed for DI and AI, changed with age and varied by society and other factors (see Supplementary Information for regression results and additional analyses). We controlled for repeated measures in all regressions.

**Data reporting.** No statistical methods were used to predetermine sample size. Within each society, we used a pseudo-random process to assign children to the two conditions to ensure a balanced sample at each age level. The investigators were not blinded to allocation during experiments and outcome assessment.

# Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota

Kristoffer Forslund<sup>1\*</sup>, Falk Hildebrand<sup>1,2,3\*</sup>, Trine Nielsen<sup>4\*</sup>, Gwen Falony<sup>2,5\*</sup>, Emmanuelle Le Chatelier<sup>6,7\*</sup>, Shinichi Sunagawa<sup>1</sup>, Edi Prifti<sup>6,7,8</sup>, Sara Vieira-Silva<sup>2,5</sup>, Valborg Gudmundsdottir<sup>9</sup>, Helle Krogh Pedersen<sup>9</sup>, Manimozhayan Arumugam<sup>4</sup>, Karsten Kristiansen<sup>10</sup>, Anita Yvonne Voigt<sup>1,11,12</sup>, Henrik Vestergaard<sup>4</sup>, Rajna Hercog<sup>1</sup>, Paul Igor Costea<sup>1</sup>, Jens Roat Kultima<sup>1</sup>, Junhua Li<sup>13</sup>, Torben Jørgensen<sup>14,15,16</sup>, Florence Levenez<sup>6,7</sup>, Joël Dore<sup>6,7</sup>, MetaHIT consortium†, H. Bjørn Nielsen<sup>9</sup>, Søren Brunak<sup>9,17</sup>, Jeroen Raes<sup>2,3,5</sup>, Torben Hansen<sup>4,18</sup>, Jun Wang<sup>10,13,19,20,21</sup>, S. Dusko Ehrlich<sup>6,7,22</sup>, Peer Bork<sup>1,12,23,24</sup> & Oluf Pedersen<sup>4</sup>

In recent years, several associations between common chronic human disorders and altered gut microbiome composition and function have been reported<sup>1,2</sup>. In most of these reports, treatment regimens were not controlled for and conclusions could thus be confounded by the effects of various drugs on the microbiota, which may obscure microbial causes, protective factors or diagnostically relevant signals. Our study addresses disease and drug signatures in the human gut microbiome of type 2 diabetes mellitus (T2D). Two previous quantitative gut metagenomics studies of T2D patients that were unstratified for treatment yielded divergent conclusions regarding its associated gut microbial dysbiosis<sup>3,4</sup>. Here we show, using 784 available human gut metagenomes, how antidiabetic medication confounds these results, and analyse in detail the effects of the most widely used antidiabetic drug metformin. We provide support for microbial mediation of the therapeutic effects of metformin through short-chain fatty acid production, as well as for potential microbiota-mediated mechanisms behind known intestinal adverse effects in the form of a relative increase in abundance of *Escherichia* species. Controlling for metformin treatment, we report a unified signature of gut microbiome shifts in T2D with a depletion of butyrate-producing taxa<sup>3,4</sup>. These in turn cause functional microbiome shifts, in part alleviated by metformin-induced changes. Overall, the present study emphasizes the need to disentangle gut microbiota signatures of specific human diseases from those of medication.

T2D is a disorder of elevated blood glucose levels (hyperglycaemia) primarily due to insulin resistance and inadequate insulin secretion, with rising global prevalence. Genetic and environmental risk factors are known, the latter including dietary habits and a sedentary lifestyle<sup>5</sup>. Gut microbiota involvement is also increasingly recognized<sup>3,4,6,7</sup>, although findings diverge between studies<sup>8</sup>; for example, Qin *et al.*<sup>3</sup> report several *Clostridium* species enriched in T2D, whereas Karlsson *et al.*<sup>4</sup> instead report enrichment of several lactobacilli species (see Supplementary Discussion). Treatment involves medication and

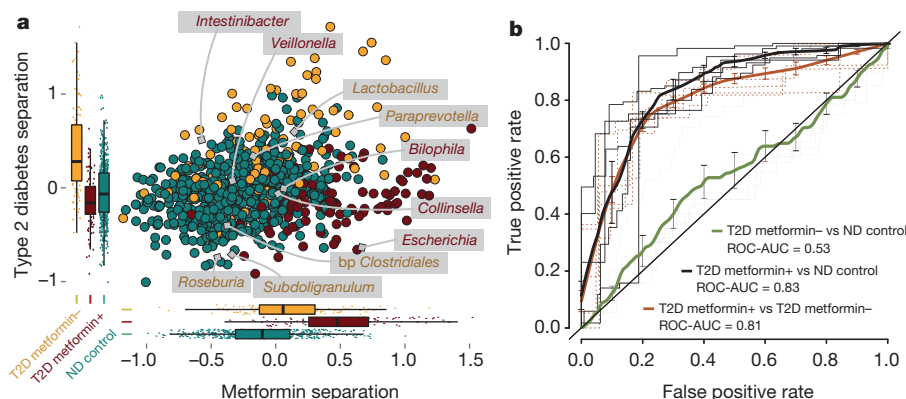
lifestyle intervention, which may confound reported gut dysbiosis. Many T2D patients receive metformin, an oral blood-glucose-lowering non-metabolizable compound whose primary and dominant metabolic effect is the inhibition of liver glucose production<sup>9</sup>. At least 30% of patients report adverse effects including diarrhoea, nausea, vomiting and bloating, with underlying mechanisms poorly understood. Studies in animals<sup>10</sup> and humans<sup>11</sup> suggest that some beneficial effects of metformin on glucose metabolism may be microbially mediated. Here, we built a multi-country T2D metagenomic data set, starting with gut microbial samples from a nondiabetic Danish cohort of 277 individuals within the MetaHIT project<sup>12</sup> and additional novel Danish MetaHIT metagenomes from 75 T2D and 31 type 1 diabetes (T1D) patients, sequenced using the same protocols (samples abbreviated as MHD). Treatment information was obtained for all MHD samples, as well as for samples from a previously reported<sup>4</sup> cohort of 53 female Swedish T2D patients, along with 92 nondiabetic individuals (43 with normal glucose tolerance, 49 with impaired glucose tolerance) (SWE) and a subgroup of 71 Chinese T2D patients with available information on antidiabetic treatment as well as 185 nondiabetic Chinese individuals<sup>3</sup> (CHN). For these 784 gut metagenomes (Supplementary Table 1), taxonomic and functional profiles were determined (see Methods), verifying our meta-analysis framework to be appropriate and robust in the context of theoretical considerations and through simulations (Supplementary Discussion 1 and Extended Data Fig. 1a), as well as characterizing differences between the data sets (Extended Data Fig. 2). Initial analysis unstratified for treatment but controlling for demographic and technical variation between data sets (Supplementary Discussion 2 and Supplementary Table 2) recovered a majority of previously reported associations (Supplementary Discussion 2 and Supplementary Table 3) but with large divergence between data sets. Suspecting confounding treatments, we tested for influence of diet and antidiabetic medications (Supplementary Discussion 3, Supplementary Table 4 and Extended Data Fig. 1b), finding an effect resulting only from use of metformin. As the fraction

<sup>1</sup>European Molecular Biology Laboratory, Structural and Computational Biology Unit, 69117 Heidelberg, Germany. <sup>2</sup>VIB Center for the Biology of Disease, Katholieke Universiteit Leuven, 3000 Leuven, Belgium. <sup>3</sup>Department of Bioscience Engineering, Vrije Universiteit Brussel, 1040 Brussels, Belgium. <sup>4</sup>The Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, 2200 Copenhagen, Denmark. <sup>5</sup>Department of Microbiology and Immunology, Rega Institute for Medical Research, Laboratory of Molecular Bacteriology, Katholieke Universiteit Leuven, 3000 Leuven, Belgium. <sup>6</sup>MICALIS, Institut National de la Recherche Agronomique, 78352 Jouy en Josas, France. <sup>7</sup>Metagenopolis, Institut National de la Recherche Agronomique, 78352 Jouy en Josas, France. <sup>8</sup>Institute of Cardiometa-bolism and Nutrition, 75013 Paris, France. <sup>9</sup>Department of Systems Biology, Center for Biological Sequence Analysis, Technical University of Denmark, 2800 Kongens Lyngby, Denmark. <sup>10</sup>Department of Biology, University of Copenhagen, 2100 Copenhagen, Denmark. <sup>11</sup>Department of Applied Tumor Biology, Institute of Pathology, University Hospital Heidelberg, 69120 Heidelberg, Germany. <sup>12</sup>Molecular Medicine Partnership Unit, University of Heidelberg and European Molecular Biology Laboratory, 69120 Heidelberg, Germany. <sup>13</sup>Beijing Genomics Institute (BGI)-Shenzhen, 518083 Shenzhen, China. <sup>14</sup>Research Centre for Prevention and Health, Capital Region of Denmark, 2600 Glostrup, Denmark. <sup>15</sup>Department of Public Health, Faculty of Health and Medical Sciences, University of Copenhagen, 2600 Copenhagen, Denmark. <sup>16</sup>Faculty of Medicine, University of Aalborg, 9100 Aalborg, Denmark. <sup>17</sup>Novo Nordisk Foundation Center for Protein Research, Disease Systems Biology, Faculty of Health and Medical Sciences, University of Copenhagen, 2200 Copenhagen, Denmark. <sup>18</sup>Faculty of Health Sciences, University of Southern Denmark, 5000 Odense, Denmark. <sup>19</sup>Princess Al Jawhara Albrahim Center of Excellence in the Research of Hereditary Disorders, King Abdulaziz University, 80205 Jeddah, Saudi Arabia. <sup>20</sup>Macau University of Science and Technology, Avenida Wai Long, Taipa, Macau, China. <sup>21</sup>Department of Medicine and State Key Laboratory of Pharmaceutical Biotechnology, University of Hong Kong, Hong Kong. <sup>22</sup>Centre for Host-Microbiome Interactions, Dental Institute Central Office, Guy's Hospital, King's College London, London SE1 9RT, UK. <sup>23</sup>Max Delbrück Centre for Molecular Medicine, 13125 Berlin, Germany. <sup>24</sup>Department of Bioinformatics, University of Würzburg, 97074 Würzburg, Germany.

\*These authors contributed equally to this work.

†A list of participants and their affiliations appears in the Supplementary Information.





**Figure 1 | Type 2 diabetes is confounded by metformin treatment.**

Major treatment effects are seen in multivariate analysis and in classifier performance. **a**, Projection of genus-level gut microbiome samples from Danish, Chinese and Swedish studies constrained by diabetic state and metformin treatment. Multivariate analysis (dbRDA plot based on Canberra distances between bacterial genera) reveals a T2D dysbiosis, which overlaps only in part with taxonomic changes in metformin-treated patients. The ordination projects all T2D metformin+ ( $n = 93$ , dark red), T2D metformin- ( $n = 106$ , orange) and ND control ( $n = 554$ , teal) gut metagenomes, with confounding country effect adjusted for. Bacterial genera that show significant effects of metformin treatment and T2D status compared to ND control, respectively (limited to top five for each), are interpolated into the plane of maximal separation based on their abundances across all samples. Marginal box/scatter plots show the separation of the constrained projection coordinates (boxes show medians/quartiles, error bars extend to most extreme value within

1.5 interquartile ranges). The T2D separation is significant (PERMANOVA  $FDR < 0.005$ ) in the joint data set and independently significant in CHN and MHD samples. The metformin separation is significant (PERMANOVA  $FDR < 0.1$ ; Canberra distances) in MHD and SWE samples. bp, butyrate-producing. **b**, Classifying T2D and metformin treatment status based on gut microbiome profiles. Support vector machine classifiers were used to separate T2D metformin+ ( $n = 93$ ), T2D metformin- ( $n = 106$ ) and ND control ( $n = 554$ ) gut metagenomes from each other based on genus-level gut microbiome taxonomic composition. Bold curves represent mean performance in hold-out testing of 1 out of 5 of the data each time, with separate tests shown as dashed curves. Error bars show  $\pm 1$  s.d. Metformin-treated T2D samples can be well separated from controls (using *Intestinibacter* abundance as the only feature), whereas distinguishing T2D metformin- samples from ND control samples works poorly even in the best case, requiring 63 distinct microbial features to achieve this separation. ROC-AUC, area under the receiving operating characteristic curve.

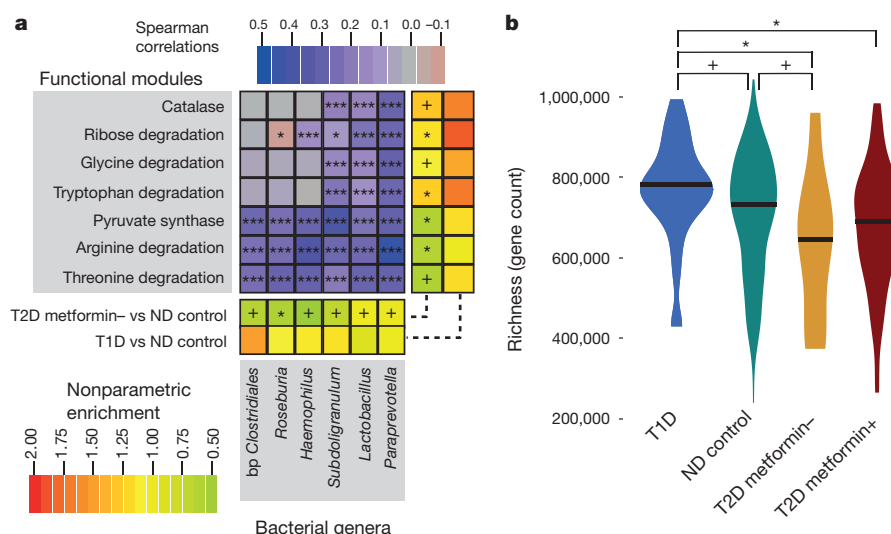
of medicated patients (denoted as T2D metformin+) varied strongly (21% CHN, 38% SWE and 77% MHD), samples were stratified on metformin treatment status. Multivariate analysis showed significant (permutational multivariate analysis of variance (PERMANOVA) false discovery rate ( $FDR$ )  $< 0.005$ ) differences in gut taxonomic composition between metformin-untreated T2D (T2D metformin-) ( $n = 106$ ) patients and nondiabetic controls (ND control) ( $n = 554$ ), consistent with a broad-range dysbiosis in T2D (Fig. 1a and Supplementary Table 5; see also Extended Data Table 1a and Supplementary Discussion 3 for an analysis of variances broken down by source). While metformin treatment status could be reliably recovered from microbial composition using support vector machines, metformin-untreated T2D status itself could not (Fig. 1b and Supplementary Table 6). In contrast, in all three cohorts, drug-treatment-blinded T2D samples could be separated from ND control samples with similar accuracy as previously reported<sup>3,4</sup>, suggesting that the T2D metformin+ classifier robustly outperforms T2D metformin- classifiers across data sets (Supplementary Table 7).

We further explored T2D gut microbiome alterations in 106 metformin-untreated T2D compared with 554 ND control samples through univariate tests of microbial taxonomic and functional differences, with significant trends shown in Fig. 2a. Metformin-untreated T2D was associated with a decrease in genera containing known butyrate producers such as *Roseburia* spp., *Subdoligranulum* spp. and a cluster of butyrate-producing Clostridiales spp. (Supplementary Table 8), consistent with previous indications<sup>3,4</sup>. More fine-grained taxonomic analysis indicated some driver species (Supplementary Discussion 4 and Supplementary Table 9), and further found changes in abundance of several unclassified Firmicutes, often reduced or reversed under metformin treatment (see Supplementary Discussion 4). Although an increase in *Lactobacillus* spp. was seen in treatment-unstratified T2D samples (as previously found experimentally<sup>13</sup>), this trend was eliminated or reversed when controlling for metformin. Functionally, we found enrichment of catalase (conceivably a response to increased peroxide stress under inflammation) and modules for ribose, glycine and tryptophan amino acid degradation,

but a decrease in threonine and arginine degradation, and in pyruvate synthase capacity (Supplementary Table 10). While these functional differences could result from strain-level composition changes or be a compound effect of subtle enrichment/depletion of larger ecological guilds, the abundance of most of these modules correlated with abundance of the significantly altered microbial genera (Fig. 2a).

To interpret our findings on T2D gut microbiota shifts further, we compared them with 31 adult T1D patients (Supplementary Table 1; for further discussion of this sub-cohort, see also Supplementary Discussion 5 and Supplementary Tables 6 and 11). This group is dysglycaemic like T2D patients, allowing us to separate purely glycaemic phenotype effects from T2D-specific microbial features. Gene richness was significantly increased in the T1D microbiomes (Wilcoxon rank sum test  $FDR < 0.1$ ) (Fig. 2b), but was reduced in T2D (Supplementary Table 10), as reported previously<sup>6</sup>. Features found to distinguish metformin-untreated T2D from ND control microbiomes did not replicate when comparing T1D to ND control. Instead, most differences between metformin-untreated T2D samples and ND controls were reversed in adult T1D patients. In contrast, some microbial functions differentially abundant between metformin-untreated T2D and controls showed similar trends in T1D samples (Fig. 2a), although not significantly, possibly owing to lower statistical power. We therefore conclude that the majority of gut microbiota shifts visible in metformin-untreated T2D are not simply effects of dysglycaemia, but rather directly or indirectly associated with the causes or progression of T2D.

Suspecting microbial mediation of some of the therapeutic effects of metformin, we next compared T2D metformin-treated ( $n = 93$ ) and T2D metformin-untreated ( $n = 106$ ) samples to characterize the treatment effect in more detail. Multivariate contrasts of T2D metformin-treated with T2D metformin-untreated samples appeared weaker than those between T2D metformin-untreated and ND control samples, the former only significant at the bacterial family level (PERMANOVA  $FDR < 0.1$ ), suggesting that the effects of metformin treatment on gut microbial composition are poorly captured by multivariate analysis. Univariate tests of the effects of metformin treatment showed a significant increase of *Escherichia* spp. and a reduced



**Figure 2 | Gut microbiome signatures in metformin-naive T2D and T1D.** Differences between healthy controls and T2D patients contrasted against T1D as an alternative form of dysglycaemia. **a**, Taxonomic and functional microbiome signatures of metformin-naive T2D. The heat maps show bacterial genera (horizontal axis) and microbial gene functions (vertical axis) that are significantly (study-source-adjusted Kruskal–Wallis test and post-hoc Mann–Whitney *U*-test, markers in innermost marginal heat maps indicating \*FDR < 0.05, +FDR < 0.1) different in abundance (nonparametric enrichment scores shown as intensity of innermost marginal heat maps; red–green colour scale) between T2D metformin– ( $n = 106$ ) and ND control ( $n = 554$ ) gut metagenomes, revealing a robust diabetic signature across data sets. None of these features is significantly

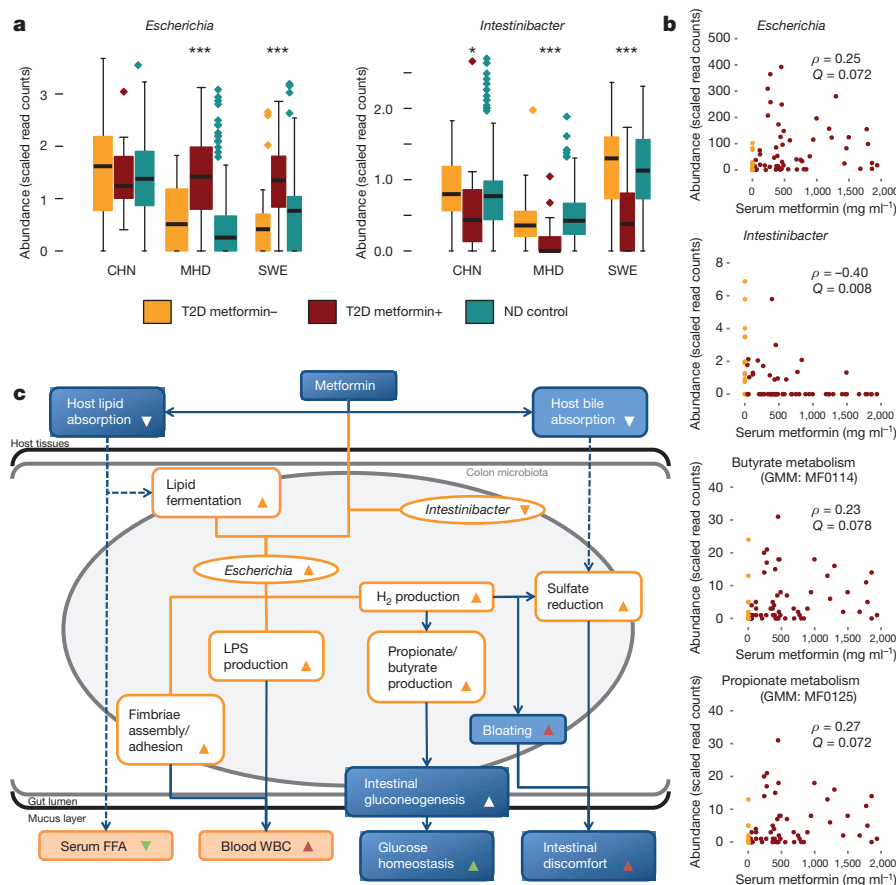
different in a comparison of T1D ( $n = 31$ ) with ND control ( $n = 277$ ) gut metagenomes (outermost marginal heat maps, same notation as above), implying that they are not direct effects of dysglycaemia. The central heat map shows Spearman correlations (purple–red colour scale) between abundance of bacterial taxa and microbial gene modules (Spearman test FDR scores: \*FDR < 0.05, \*\*\*FDR < 0.001). **b**, Elevated gene richness in adult type 1 diabetes samples. Comparing MHD samples only, T1D ( $n = 31$ ) gut metagenomes show significantly (Mann–Whitney *U*-test, +FDR < 0.1, \*FDR < 0.05) higher gut microbiome richness (that is, gene count) than all other sample subsets ( $n = 277$  ND control,  $n = 58$  T2D metformin–,  $n = 17$  T2D metformin+ gut metagenomes). Sample median richness is shown as horizontal black bars.

abundance of *Intestinibacter* spp., the latter fully consistent across the different country data sets (Fig. 3a), whereas the former is not seen in the CHN cohort where both diabetic individuals and controls are enriched in *Escherichia* spp. relative to Scandinavian controls. Correcting for differences in gender, body mass index and fasting levels of plasma glucose or serum insulin (some of which were significantly different between data sets, Supplementary Table 12) retained these differences as significant (Supplementary Table 13). Fasting serum concentrations of metformin were obtained for the MHD cohort and correlated significantly with abundances of both genera (Fig. 3b). Amplicon-based analysis of an independent T2D cohort likewise validated an increase of *Escherichia* spp. and a reduced abundance of *Intestinibacter* spp. in metformin-treated patients (Extended Data Fig. 1c, Extended Data Table 1b and Supplementary Discussion 6). The metformin-associated changes might derive from taxon-specific resistance/sensitivity to the bacteriostatic or bactericidal properties of the drug<sup>14</sup>. The genus *Intestinibacter* was defined only recently<sup>15</sup> and includes the human isolate *Clostridium bartlettii*<sup>16</sup>, since reclassified as *Intestinibacter bartlettii*. Little is known about its role in the gut ecosystem and how it might affect human health. However, *I. bartlettii* abundances were lower in pigs susceptible to colonization by enterotoxigenic *Escherichia* spp.<sup>17</sup>, consistent with the pattern seen here following metformin treatment. Analysis of the SEED (see Supplementary Discussion 7) and GMM (see Methods) functional annotations linked to *Intestinibacter* shows it to be resistant to oxidative stress and able to degrade fucose, indicative of an indirect involvement in mucus degradation. It also appears to possess the genetic potential for sulfite reduction, including part of an assimilatory sulfate reduction pathway. Analysis of gut microbial functional potential more generally suggested that indirect metformin treatment effects (Fig. 3c), including reduced intestinal lipid absorption<sup>18</sup> and lipopolysaccharide (LPS)-triggered local inflammation, can provide a competitive advantage to *Escherichia* species<sup>19</sup>, possibly triggering a positive feedback loop that further contributes to the observed taxonomic changes. At the same time, metformin may reverse T2D-associated changes, as

several gut microbial genera were more similar in abundance to ND control levels under metformin treatment, notably *Subdoligranulum* and to some extent *Akkermansia*. The latter was previously shown to reduce insulin resistance in murine models when increased in abundance through prebiotics<sup>20</sup>, and has been shown to similarly increase in abundance under metformin treatment<sup>10,21</sup>. In human samples, however, the trend was inconsistent between country subsets, and only MHD samples show a similar response (Extended Data Fig. 3). With respect to microbiota-mediated impact on host glucose regulation, the functional analyses demonstrated significantly enhanced butyrate and propionate production potential in metformin-treated individuals (Fig. 3c and Supplementary Table 14). Interestingly, recent studies in mice have shown that an increase in colonic production of these short-chain fatty acids triggers intestinal gluconeogenesis (IGN) via complementary mechanisms. Butyrate activates IGN gene expression through a cAMP-dependent mechanism in enterocytes, whereas propionate, itself a substrate of IGN, activates IGN gene expression via the portal nervous system and the fatty acid receptor FFAR3 (refs 22, 23). In rodents, the net result of increased IGN is a beneficial effect on glucose and energy homeostasis with reductions in hepatic glucose production, appetite and body weight. Taken together, our characterization of a metformin-associated human gut microbiome suggests novel mechanisms contributing to the beneficial effects of the drug on host metabolism.

Both on a compositional and functional level, we found significant microbiome alterations that are consistent with well-known side-effects of metformin treatment (Fig. 3c). Most of these metformin-associated functional shifts, including enrichment of virulence factors and gas metabolism genes, could be attributed to the significantly increased abundance of *Escherichia* species (Supplementary Discussion 7 and Supplementary Tables 14 and 15).

In conclusion, our results suggest partial gut microbial mediation of both therapeutic and adverse effects of the most widely used antidiabetic medication, metformin, although further validation is required to conclude causality and to clarify how such mediation might occur.



**Figure 3 | Impact of metformin on the human gut microbiome.**

Characterization of the microbially mediated therapeutic and adverse effects of metformin. **a**, Gut microbial shifts under metformin treatment. Metformin treatment significantly (study-source-adjusted Kruskal–Wallis test and post-hoc Mann–Whitney *U*-test, \*FDR < 0.05, \*\*\*FDR < 0.001) increases *Escherichia* and lowers *Intestinibacter* abundance. Box plots show median/quartile abundances, whiskers extend to  $1.58 \times$  interquartile range/ $\sqrt{n}$ , for T2D metformin+ ( $n_{\text{CHN}} = 15$ ,  $n_{\text{MHD}} = 58$ ,  $n_{\text{SWE}} = 20$ ), T2D metformin- ( $n_{\text{CHN}} = 56$ ,  $n_{\text{MHD}} = 17$ ,  $n_{\text{SWE}} = 33$ ) and ND control ( $n_{\text{CHN}} = 185$ ,  $n_{\text{MHD}} = 277$ ,  $n_{\text{SWE}} = 92$ ) gut metagenome samples. **b**, Correlations between serum levels of metformin and gut microbiota in Danish MetaHIT samples, including short-chain fatty acid production modules. Serum metformin levels of T2D patients ( $n = 75$  gut metagenomes) are significantly (Spearman FDR < 0.1) positively correlated with *Escherichia* abundance, and in significant negative correlation with *Intestinibacter* abundance. Bacterial gene function modules for butyrate and propionate

production increase in abundance as serum metformin levels increase.

Dot markers are shown for all MHD samples for which serum metformin concentration was measured. Metformin-untreated T2D samples (serum concentration < 10 mg ml<sup>-1</sup>) are shown in orange, treated samples in dark red. Spearman coefficients ( $\rho$ ; calculated for treated samples only) and FDRs (*Q*) are shown. **c**, Microbial shifts under metformin treatment contribute to improved glucose control and to adverse effects. Schematic illustration of gut microbial changes and their impact on host health. Observed associations (orange lines) between microbial taxa abundances (orange ellipses), microbial functional potential (orange boxes), and blood values (filled orange boxes) and metformin treatment are linked with literature-derived metformin- or microbiota-induced host physiological effects (blue boxes and arrows; dashed arrows indicate hypothesized causality). Drug–host–microbiota interactions can contribute to previously described therapeutic (green triangles) and side (red triangles) effects of metformin treatment.

Our study of T2D illustrates the need to disentangle specific disease dysbioses from effects of treatment on the human-associated microbiota. The importance of this point was further shown by the fact that the previously reported high accuracy<sup>3,4</sup> of gut microbial signatures for identifying patients with treatment-unstratified T2D decreased markedly when considering a large set of metformin-naïve patients only, highlighting a general need to bear treatment regimens in mind both when developing and applying microbiome-based diagnostic and prognostic tools for common disorders or their pre-morbidity states.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 4 March; accepted 5 October 2015.

Published online 2 December 2015.

- Shreiner, A. B., Kao, J. Y. & Young, V. B. The gut microbiome in health and in disease. *Curr. Opin. Gastroenterol.* **31**, 69–75 (2015).
- Cho, I. & Blaser, M. J. The human microbiome: at the interface of health and disease. *Nature Rev. Genet.* **13**, 260–270 (2012).

- Qin, J. *et al.* A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490**, 55–60 (2012).
- Karlsson, F. H. *et al.* Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* **498**, 99–103 (2013).
- Schellenberg, E. S., Dryden, D. M., Vandermeer, B., Ha, C. & Korownyk, C. Lifestyle interventions for patients with and at risk for type 2 diabetes: a systematic review and meta-analysis. *Ann. Intern. Med.* **159**, 543–551 (2013).
- Larsen, N. *et al.* Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. *PLoS ONE* **5**, e9085 (2010).
- Zhang, X. *et al.* Human gut microbiota changes reveal the progression of glucose intolerance. *PLoS ONE* **8**, e71108 (2013).
- de Vos, W. M. & Nieuwdorp, M. Genomics: A gut prediction. *Nature* **498**, 48–49 (2013).
- Pernicova, I. & Korbonits, M. Metformin—mode of action and clinical implications for diabetes and cancer. *Nat. Rev. Endocrinol.* **10**, 143–156 (2014).
- Shin, N. R. *et al.* An increase in the *Akkermansia* spp. population induced by metformin treatment improves glucose homeostasis in diet-induced obese mice. *Gut* **63**, 727–735 (2014).
- Napolitano, A. *et al.* Novel gut-based pharmacology of metformin in patients with type 2 diabetes mellitus. *PLoS ONE* **9**, e100778 (2014).
- Le Chatelier, E. *et al.* Richness of human gut microbiome correlates with metabolic markers. *Nature* **500**, 541–546 (2013).
- Sato, J. *et al.* Gut dysbiosis and detection of “live gut bacteria” in blood of Japanese patients with type 2 diabetes. *Diabetes Care* **37**, 2343–2350 (2014).



14. Cabreiro, F. *et al.* Metformin retards aging in *C. elegans* by altering microbial folate and methionine metabolism. *Cell* **153**, 228–239 (2013).
15. Gerritsen, J. *et al.* Characterization of *Romboutsia ilealis* gen. nov., sp. nov., isolated from the gastro-intestinal tract of a rat, and proposal for the reclassification of five closely related members of the genus *Clostridium* into the genera *Romboutsia* gen. nov., *Intestinibacter* gen. nov., *Terrisporobacter* gen. nov. and *Asaccharospora* gen. nov. *Int. J. Syst. Evol. Microbiol.* **64**, 1600–1616 (2014).
16. Song, Y. L., Liu, C. X., McTeague, M., Summanen, P. & Finegold, S. M. *Clostridium bartlettii* sp. nov., isolated from human faeces. *Anaerobe* **10**, 179–184 (2004).
17. Messori, S., Trevisi, P., Simongiovanni, A., Priori, D. & Bosi, P. Effect of susceptibility to enterotoxigenic *Escherichia coli* F4 and of dietary tryptophan on gut microbiota diversity observed in healthy young pigs. *Vet. Microbiol.* **162**, 173–179 (2013).
18. Czyzyk, A., Taweccki, J., Sadowski, J., Ponikowska, I. & Szczepanik, Z. Effect of biguanides on intestinal absorption of glucose. *Diabetes* **17**, 492–498 (1968).
19. Winter, S. E. *et al.* Host-derived nitrate boosts growth of *E. coli* in the inflamed gut. *Science* **339**, 708–711 (2013).
20. Everard, A. *et al.* Cross-talk between *Akkermansia muciniphila* and intestinal epithelium controls diet-induced obesity. *Proc. Natl Acad. Sci. USA* **110**, 9066–9071 (2013).
21. Lee, H. & Ko, G. Effect of metformin on metabolic improvement and gut microbiota. *Appl. Environ. Microbiol.* **80**, 5935–5943 (2014).
22. De Vadder, F. *et al.* Microbiota-generated metabolites promote metabolic benefits via gut-brain neural circuits. *Cell* **156**, 84–96 (2014).
23. Croset, M. *et al.* Rat small intestine is an insulin-sensitive gluconeogenic organ. *Diabetes* **50**, 740–746 (2001).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** The authors wish to thank A. Forman, T. Lorentzen, B. Andreassen, G. J. Klavsén and M. J. Nielsen for technical assistance, and T. F. Toldsted and G. Lademann for management assistance. J. Nielsen and F. Bäckhed are thanked for providing access to T2D metagenome data and metformin treatment status before publication<sup>4</sup>. V. Benes and the GeneCore facility of EMBL Heidelberg are thanked for their assistance with the metformin signature validation experiments, as is Y. Yuan for assistance with computer infrastructure. This research has received funding from European Community's Seventh Framework Program (FP7/2007–2013): MetaHIT, grant agreement HEALTH-F4-2007-201052, MetaCardis, grant

agreement HEALTH-2012-305312, International Human Microbiome Standards, grant agreement HEALTH-2010-261376, as well as from the Metagenopolis grant ANR-11-DPBS-0001, from the European Research Council CancerBiome project, contract number 268985, and from the European Union HORIZON 2020 programme, under Marie Skłodowska-Curie grant agreement 600375. Additional funding came from The Lundbeck Foundation Centre for Applied Medical Genomics in Personalized Disease Prediction, Prevention and Care (LuCamp, <http://www.lucamp.org>), the Novo Nordisk Foundation (grant NNF14CC0001), and the European Molecular Biology Laboratory (EMBL). The Novo Nordisk Foundation Center for Basic Metabolic Research is an independent Research Center at the University of Copenhagen partially funded by an unrestricted donation from the Novo Nordisk Foundation (<http://www.metabol.ku.dk>). Additional funding for the validation experiments was provided by the Innovation Fund Denmark through the MicrobDiab project.

**Author Contributions** O.P., S.D.E. and P.B. devised the project, designed the study protocol and supervised all phases of the project. T.N., T.H., T.J., H.V., J.L. and O.P. carried out patient phenotyping and clinical data analyses. T.N. and F.L. performed sample collection and DNA extraction. J.D. supervised DNA extraction, J.W., K.K. supervised DNA sequencing and gene profiling, A.Y.V. and R.H. performed additional microbial DNA extraction and amplicon sequencing. J.R., H.B.N., S.B., S.D.E., P.B. and O.P. designed and supervised the data analyses. K.F., F.H., G.F., E.L.C., S.S., E.P., S.S.-V., V.G., H.K.P., M.A., P.I.C., J.R.K. and H.B.N. performed the data analyses. K.F., F.H., T.N., P.B., S.D.E. and O.P. wrote the paper. All authors contributed to data interpretation, discussions and editing of the paper. All authors are members of the MetaHIT consortium. Additional consortium members contributed to the design and execution of the study.

**Author Information** Raw nucleotide data can be found for all samples used in the study in the Sequence Read Archive (accession numbers: SRA045646 and SRA050230, CHN samples) and the European Nucleotide Archive (accession numbers: ERP002469, SWE samples; ERA000116, ERP003612, ERP002061 and ERP004605, MHD samples). Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.D.E. ([dusko.ehrlich@jouy.inra.fr](mailto:dusko.ehrlich@jouy.inra.fr)) or P.B. ([bork@embl.de](mailto:bork@embl.de)) or O.P. ([oluf@sund.ku.dk](mailto:oluf@sund.ku.dk)).

## METHODS

No statistical methods were used to predetermine sample size.

**Danish MetaHIT diabetic study.** *Patient recruitment, enrolment and processing.* Patients with T2D were either recruited from the Inter99 study population<sup>24</sup> or from the out-patient clinic at Steno Diabetes Center, Gentofte, Denmark. Patients with known T2D were included if the patient had clinically defined T2D on the day of examination according to the WHO definition<sup>25</sup>. Inclusion criteria were fasting serum C-peptide above 200 pmol l<sup>-1</sup> and negative testing for serum glutamic acid decarboxylase (GAD) 65 antibodies (to exclude T1D, latent autoimmune diabetes in adults), no secondary forms of diabetes like chronic pancreatitis diabetes or syndromic diabetes, no antibiotic treatment 2 months before inclusion, and no known gastro-intestinal diseases, no previous bariatric surgery or medication known to affect the immune system.

All patients with T1D were recruited from the out-patient clinic at Steno Diabetes Center, Gentofte, Denmark ( $n = 31$ ). Inclusion criteria were dependence on insulin treatment from time of diagnosis, fasting serum C-peptide below 200 pmol l<sup>-1</sup>, glycated haemoglobin (HbA1c) above 8.0% (64 mmol l<sup>-1</sup>) to ensure current hyperglycaemia, T1D duration and dependence on insulin treatment > 5 years, no antibiotic treatment at least 2 months before inclusion, and no known gastrointestinal diseases. All study participants were of North European ethnicity.

The study participants were examined on 2 days that were approximately 14 days apart. On the first day, study participants were examined after an over-night fast. Height was measured without shoes to the nearest 0.5 cm, and weight was measured without shoes and wearing light clothes to the nearest 0.1 kg. Hip and waist circumference was measured using a non-expandable measuring tape to the nearest 0.5 cm. Waist circumference was measured midway between the lower rib margin and the iliac crest. Hip circumference was measured as the largest circumference between the waist and the thighs. Blood pressure was assessed while the participant was lying in an up-right position after at least 5 min of rest using a cuff of appropriate size (A&D, UA-787 plus digital or A&D, UA-779). Blood pressure was measured at least twice and the average of the measurements was calculated. On the second day of examination, all participants provided a stool sample which was immediately frozen after home collection and stored at  $-80^{\circ}\text{C}$ .

Information on medication status was obtained by questionnaire and interview on the first day of examination. Of the 75 T2D patients, 10 patients (13%) received no hyperglycaemic medications and 58 patients (77%) received the biguanide metformin; of these 75 T2D patients, 28 patients (37%) received metformin as the only anti-hyperglycaemic medication, 10 patients (13%) received sulfonylurea alone or in combination with metformin, 14 patients (19%) received a combination of oral antidiabetic drugs and insulin treatment and 4 patients (5%) were on insulin treatment only. Eleven patients (15%) received dipeptidyl peptidase-4 (DPP4) inhibitors or glucagon-like peptide-1 (GLP1), all of them in combination with metformin. Patients were reported as receiving anti-hypertensive treatment if at least one of the following drugs was reported: spironolactone, thiazides, loop diuretics, beta blockers, calcium channel blockers, moxonidine or drugs affecting the renin-angiotensin system ( $n = 55$  for T2D (73%) and  $n = 23$  (74%) for T1D). Patients receiving statins, fibrates and/or ezetimibe were reported as receiving lipid-lowering medication ( $n = 56$  for T2D (75%; all on statin treatment), and  $n = 24$  for T1D (77%; 74% on statin treatment)). All T1D patients were on insulin treatment as their only blood glucose lowering treatment.

All biochemical analyses were performed on blood samples drawn in the morning after an over-night fast of at least 10 h. Plasma glucose was analysed by a glucose oxidase method (Granustest, Merck) with a detection limit of 0.11 mmol l<sup>-1</sup> and intra- and interassay coefficients of variation (CV) of <0.8% and <1.4%, respectively. HbA1c was measured on G7 HPLC Analyzer (Tosoh) by ion-exchange high-performance liquid chromatography. Serum C-peptide was measured using a time-resolved fluoroimmunoassay with the AutoDELFIA C-peptide kit (PerkinElmer, Wallac), with a detection limit of 5 pmol l<sup>-1</sup> and intra- and interassay CV of <4.7% and <6.4%, respectively. Serum insulin (excluding des and intact proinsulin) was measured using the AutoDELFIA insulin kit (PerkinElmer, Wallac) with a detection limit of 3 pmol l<sup>-1</sup> and with intra- and interassay CV of <3.2% and <4.5%, respectively. Plasma cholesterol, plasma high-density lipoprotein cholesterol and plasma triglycerides were all measured on Vitros 5600 using reflect-spectrophotometrics. Plasma low-density lipoprotein cholesterol was calculated using Friedewald's equation. Blood leukocytes and white blood cell differential count were measured on Sysmex XS 1000i using flow cytometrics. Plasma metformin was determined by high performance liquid chromatography followed by tandem mass spectrometry. Briefly, the proteins were precipitated

with acetonitrile containing the deuterated internal standard, metformin-d6, hydrochloride and the supernatant diluted by acetonitrile. The analysis was performed on a Waters Acquity UPLC I-class system connected to a Xevo TQ-S tandem mass spectrometer in electrospray positive ionization mode. Separation was achieved on a Waters XBridgeT BEH Amide 2.5- $\mu\text{m}$  column and gradient elution with 100 mM ammonium formate (pH 3.2), and with acetonitrile. The multiple reaction monitoring transitions used for metformin and metformin-d6 were 130.2 > 71.0 and 136.2 > 60.0. Calibrators were prepared by spiking drug-free serum with metformin to a concentration of 2,000 ng ml<sup>-1</sup>. B12 was measured using Vitros Immunodiagnostic Products. GAD65 was measured on serum samples by a sandwich ELISA (RSR Ltd.). Inter- and intra-assay CV were < 16.6% and < 6.7% respectively, and with a detection limit of 0.57 U ml<sup>-1</sup>.

Stool samples were obtained at the homes of each participant and samples were immediately frozen by storing them in their home freezer. Frozen samples were delivered to Steno Diabetes Center using insulating polystyrene foam containers, and then they were stored at  $-80^{\circ}\text{C}$  until analysis. The time span from sampling to delivery at the Steno Diabetes Center was intended to be as short as possible and no more than 48 h.

A frozen aliquot (200 mg) of each faecal sample was suspended in 250  $\mu\text{l}$  of guanidine thiocyanate, 0.1 M Tris, pH 7.5, and 40  $\mu\text{l}$  of 10% *N*-lauroylsarcosine. Microbial DNA extraction was then performed as previously described<sup>12</sup>. The DNA concentration and its molecular size were estimated using nanodrop (Thermo Scientific) and agarose gel electrophoresis.

**Generation and availability of metagenomic samples.** Already available Danish metagenomic samples were those reported in ref. 26 and references therein (excluding 14 samples removed due to average read length below 40 nucleotides, and with 5 Chinese and 21 Swedish samples with less than the rarefaction threshold of 7 million reads in total excluded from functional profile or diversity analyses), with newly sequenced samples deposited in the European Bioinformatics Institute Sequence Read Archive under accession ERP004605.

All information on Swedish samples was retrieved from previously published data<sup>4</sup>. In addition to published data on Chinese individuals<sup>3</sup>, we retrieved information on metformin treatment in a subset of 71 Chinese T2D patients. One-hundred and twelve samples from ref. 3 lacked metformin treatment metadata and were therefore discarded, except for measuring differences between the country data sets disregarding treatment or diabetic status. Characteristics of all study participants included in the present protocol are given in Supplementary Table 1.

**Validation cohort recruitment and sample processing.** Additional Danish T2D patients were recruited at the Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen throughout 2014 as a part of the ongoing MicrobDiab study (<http://metabol.ku.dk/research-project-sites/microbdiab/>). T2D patients were included in the study if the time of T2D diagnosis was less than 5 years ago, they were between 35 and 75 years of age, Caucasian and they had not received antibiotics within the past 4 months of inclusion. In total, 30 T2D patients (21 male and 9 female) were identified. Faecal samples were collected at the home of the patients, followed by immediate freezing of samples in home freezers, and transport of samples to the hospital stored on dry ice. The samples were stored at  $-80^{\circ}\text{C}$  until DNA extraction. Information of medication was obtained from questionnaires. In total, 21 (70%) of the T2D patients received metformin.

**Ethics statement.** All individuals in both the Danish MetaHIT study and the Danish validation study gave written informed consent before participation in the studies. Both studies were approved by the Ethical Committees of the Capital Region of Denmark (MetaHIT study: HC-2008-017; validation study: H-3-2013-102). Both studies were conducted in accordance with the principles of the Declaration of Helsinki.

**Construction of a non-redundant metagenomic reference gene catalogue.** Illumina shotgun sequencing was applied to DNA extracted from 620 faecal samples originating from the MetaHIT project (Supplementary Table 1). Raw sequencing data were processed using the MOCAT (version 1.1) software package<sup>27</sup>. Reads were trimmed (option `read_trim_filter`) using a quality and length cut-off of 20 and 30 bp, respectively. Trimmed reads were subsequently screened against a custom database of Illumina adapters (option `screen_fastfile`) and the human genome version 19 using a 90% identity cut-off (option `screen`). The resulting high-quality reads were assembled (option `assembly`) and assemblies revised (option `assembly_revision`). Genes were predicted on scaffolds with a minimum length of 500 bp (option `gene_prediction`).

Predicted protein-coding genes with a minimum length of 100 bp were clustered at 95% sequence identity using Cd-hit (version 4.6.1)<sup>28</sup> with parameters set to: `-c 0.95`, `-G 0`, `-a 0.9`, `-g 1`, `-r 1`. The representative genes of the resulting clusters were 'padded' (that is, extended up to 100 bp at each end of the sequence using

the sequence information available from the assembled scaffolds), resulting in the final reference gene catalogue used in this study.

The reference gene catalogue was functionally annotated using SmashCommunity<sup>29</sup> (version 1.6) after aligning the amino acid sequence of each gene to the KEGG<sup>30</sup> (version 62) and eggNOG<sup>31</sup> (version 3) databases.

**Profiling of metagenomic samples.** Raw insert (sequenced fragments of DNA represented by single or paired-end reads) count profiles were generated using MOCAT<sup>27</sup> by mapping high-quality reads from each metagenome to the reference gene catalogue (option screen) using an alignment length and identity cut-off of 45% and 95%, respectively. For each gene, the number of inserts that matched the protein-coding region was counted. Counts of inserts that mapped with the same alignment score to multiple genes were distributed equally among them. Taxonomic abundances were computed at the level of metagenomic operational taxonomic units (mOTUs)<sup>32</sup>, normalized to the length of the concatenated marker genes for each mOTU to yield the abundances used for the study, and subsequently binned at broader taxonomic levels (genus, family, class, etc.).

**Rarefaction of metagenomic data and microbial diversity measurements.** For all metagenome-derived measures except the mOTU taxonomic assignments, read counts were 'rarefied' in order to avoid any artefacts of sample size on low-abundance genes. Rarefied matrices were obtained as follows. Data matrices were rarefied to 7 million reads per sample. This threshold was chosen to include most samples, but 5 Chinese and 21 Swedish samples were excluded due to having less than 7 million reads per sample. Rarefactions were performed using a C++ program developed for the Tara project<sup>33</sup>. In total we performed 30 repetitions, and in each of these we measured the richness, evenness, chao1 and Shannon diversity metrics within a rarefaction. The median value of these was taken as the respective diversity measurement for each sample. The first of 30 rarefactions of each sample were used to create a rarefied gene abundance matrix and KEGG orthologue abundance profiles were calculated by summing the rarefied abundance of genes annotated to the respective KEGG orthologue gene.

**Metagenomic species (MGS) construction.** Clustering of the catalogue genes by co-abundance, as described in ref. 34, defined 10,754 co-abundance gene groups (CAGs) with very high correlations (Pearson correlation coefficient > 0.9). The 925 largest of these, with more than 700 genes, were termed metagenomic species (MGS). The abundance profiles of the CAGs and MGSs were determined as the medium gene abundance (downsized to 7 million reads per sample) throughout the samples. Furthermore, the CAGs and MGS were taxonomically annotated by sequence similarity to known reference genomes.

**Functional annotation/binning of metagenomes.** To avoid drawing false conclusions about gut microbial functions from high abundance of single genes remotely homologous to members of a functional pathway, we used an approach that required presence of multiple pathway members. Functional pathway abundance was calculated from gene catalogue KEGG orthologue annotation and MGS abundances per sample. Thus KEGG orthologues present in each MGS were used to determine for that CAG/MGS which functional modules were represented within its genetic repertoire. This required that >90% of KEGG orthologues necessary for the completion of a reaction pathway should be present, when also taking alternative enzymatic pathways into account. The module abundance within a sample was calculated from CAG abundance in each respective sample, summing over all CAGs which had the module present. Rarefied median coverages of CAG/MGS were used, so no further normalization of the module abundance matrix was required. Abundance of genetic potential falling under the same higher-order functional levels was calculated by summing up all abundances of the lower-level functional modules within each sample.

Existing functional annotation databases cover gut metabolic pathways relatively poorly. To account for this, a number of additional bacterial gene functional modules were curated and annotated, extending the KEGG system; these are referred to in result tables as GMMs (gut microbial modules) and were previously described in ref. 12.

**16S amplicon processing.** 16S amplicons from frozen samples were sequenced 300 bp and 200 bp paired-end reads using an Illumina miSeq machine. We used the LotuS<sup>35</sup> pipeline in short amplicon mode with default quality filtering, clustering and denoising operational taxonomic units (OTUs) with UPARSE<sup>36</sup>, removing chimaeric OTUs against the RDP reference database ([http://drive5.com/uchime/rdp\\_gold.fa](http://drive5.com/uchime/rdp_gold.fa)) with uchime<sup>37</sup>, merging reads with FLASH<sup>38</sup> and assigning a taxonomy against the SILVA 119 rRNA database<sup>39</sup>, and further refined by BLAST searches against the NCBI rRNA database<sup>40</sup> to identify *Intestinibacter* OTUs, using the following LotuS command line options: '-p miSeq -refDB SLV -doBlast blast -amplicon\_type SSU -tax\_group bacteria -derepMin 2 -CL 2 -thr 14'.

**Univariate tests of taxonomic or functional abundance differences.** Microbial taxa where mean abundance over all samples was less than 30 reads, or that were

present in less than 3 samples, were excluded from univariate and classifier analyses. All abundances were normalized by total sample sum. For module tables, no feature filters were used except requiring the module to be present in at least 20 samples. Filtered data tables were made available online (<http://vm-lux.embl.de/~forslund/t2d/>).

Univariate testing for differential abundances of each taxonomic unit between two or more groups was tested using Mann-Whitney-U or Kruskal-Wallis tests, respectively, corrected for multiple testing using the Benjamini-Hochberg false discovery rate control procedure (Q values)<sup>41</sup>. Post-hoc statistical testing for significant differences between all combinations of two groups was conducted only for taxa with abundances significantly different at  $P < 0.2$ . Wilcoxon rank-sum tests were calculated for all possible group combinations and corrected for multiple testing again using the Benjamini-Hochberg false discovery rate, as implemented in R. When controlling for potential confounders such as source study, we used blocked 'independence\_test' function calls with options 'ytrafo = rank, teststat = scalar' for blocked WRST and 'ytrafo = rank, teststat = quad' for blocked Kruskal-Wallis test, as implemented in the COIN software package<sup>42</sup> for R. Similarly, we applied these independence tests in the framework of post-hoc testing as described above.

Analysis of correlations between taxonomic or functional features, community diversity indices and sample metadata variables were conducted using Spearman correlation tests as implemented in R, and corrected for multiple testing using the Benjamini-Hochberg false discovery rate control procedure. To control for confounders such as source study in univariate correlation analyses, blocked Spearman tests as implemented in COIN (settings 'independence\_test', options 'ytrafo = rank, xtrafo = rank, distribution = asymptotic') were used.

In some analyses, taxa were corrected for the influence of a continuous confounder variable such as microbial community richness; in these cases, the residual of a linear model between normalized log-transformed taxa abundances and overall sample gene richness was used to correct for the confounding variable. Power analysis was conducted by randomly subsampling to a given sample number, repeated 5 times to achieve robust results.

**Ordinations and multivariate tests.** All ordinations (NMDS, dbRDA) and subsequent statistical analyses were calculated using the R package vegan<sup>43</sup> using Canberra distances on normalized taxa abundance matrices, then visualized using the ggplot2 R package<sup>44</sup>. Community differences were calculated using a permutation test on the respective NMDS reduced feature space, as implemented in vegan.

Furthermore, we calculated intergroup differences for the microbiota using PERMANOVA<sup>45</sup> as implemented in vegan. This test compares the intragroup distances to the intergroup distances in a permutation scheme and from this calculates a  $P$  value. For all PERMANOVA tests, we used  $2 \times 10^5$  randomizations and a normalized genus-level mOTU abundance matrix, using Canberra intersample distances. PERMANOVA post-hoc  $P$  values were corrected for multiple testing using the Benjamini-Hochberg false discovery rate control procedure. Analysis of variance broken down by cohort, treatment and disease status was conducted by fitting these distances to a linear model of sample metadata distances, as further described in Supplementary Discussion 3.2.

**Classifier construction and evaluation.** To create classifiers for separating samples from different subsets, an L1 restricted LASSO using the R glmnet package<sup>46</sup> was carried out to test for an optimal value of lambda (number of features to be used in the final predictor) in a fivefold cross-validated and internally fourfold cross-validated LASSO run on all data. After this, the previously determined value of lambda was manually controlled for number of features used against the root mean square error of the classifier. In a fivefold cross-validation, an independent LASSO classifier was trained on 4/5 of the data using the previously determined value of lambda, and response values were predicted on 1/5 of the data. LASSO models with a Poisson response type were used in all cases.

Binary classifications between T2D and ND control samples were performed with an R reimplementation of the robust recursive feature elimination support vector machine (rrFE-SVM)<sup>47</sup> procedure. The SVM was performed in an outer cross-validation scheme on 4/5 of the data. Of these, 90% were randomly selected 200 times in each cross-validation for the RFE, to create a feature ranking from an average over these runs. Classifier performance was validated on the remaining 1/5 of samples using the pre-established feature ranking. In case of several cohorts, the area under the receiver operating characteristic curve (ROC-AUC) scores were measured for each cohort separately.

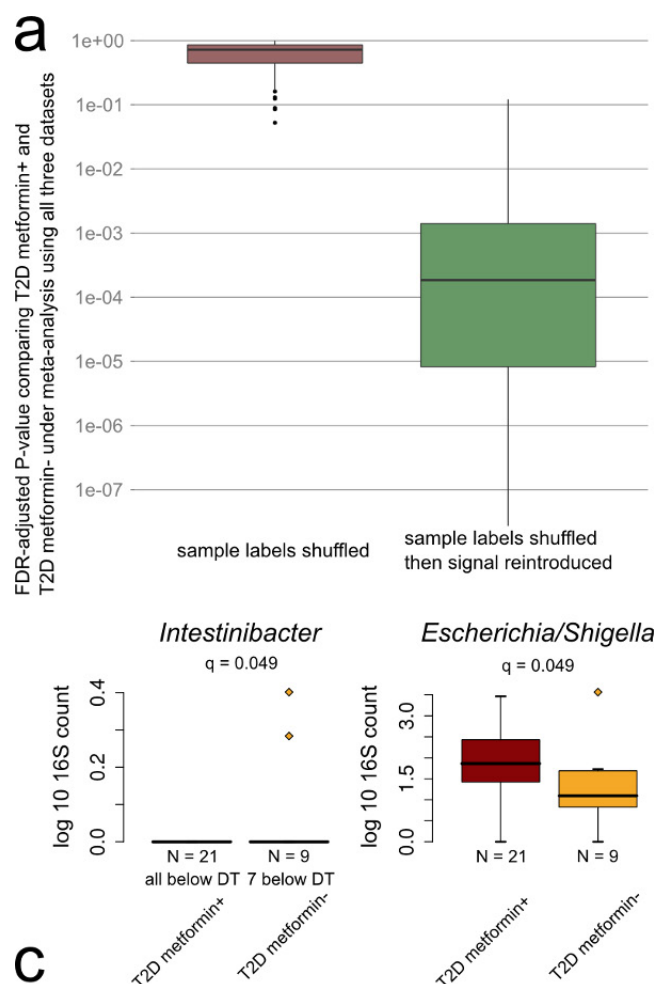
**Code availability.** The MGS technology has previously been described<sup>34</sup> and is available online (<http://git.dworzynski.eu/mgs-canopy-algorithm/wiki/Home>). The mOTU resource has been made publically available (<http://www.bork.embl.de/software/mOTU/>) and was analysed using MOCAT<sup>27</sup> which is also publically



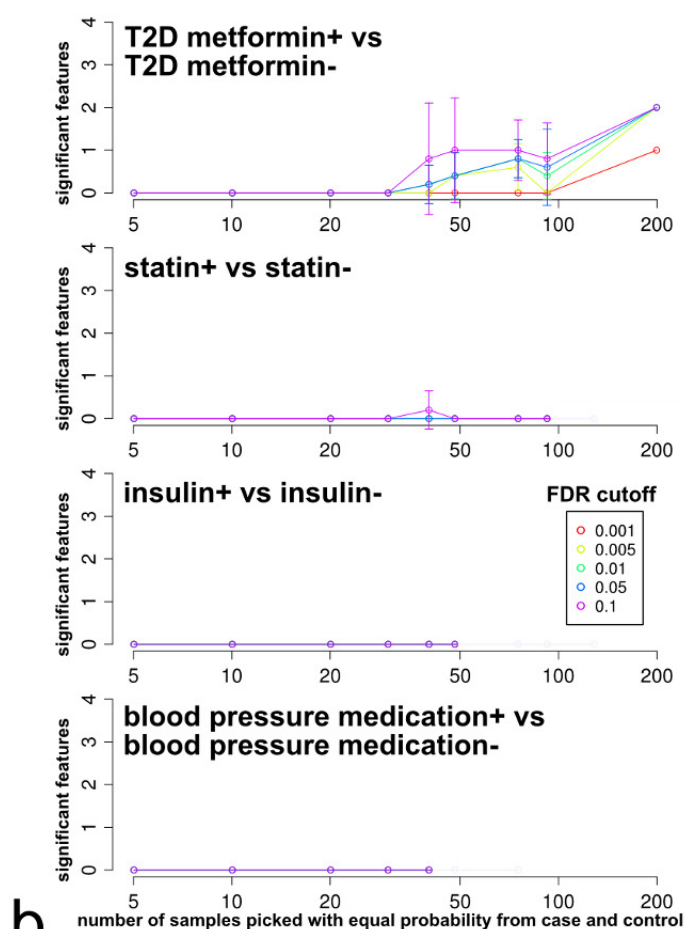
available (<http://vm-lux.embl.de/~kultima/MOCAT/>). The 16S pipeline LotuS<sup>35</sup> is freely available online (<http://psbweb05.psb.ugent.be/lotus/>). The novel gene catalogue has been deposited online ([http://vm-lux.embl.de/~kultima/share/gene\\_catalogs/620mhT2D/](http://vm-lux.embl.de/~kultima/share/gene_catalogs/620mhT2D/)), as have the raw amplicon sequences (<http://vm-lux.embl.de/~forslund/t2d/>). Statistical analysis and data visualization was conducted using freely available R libraries: vegan, COIN and ggplot2 and is described in more details elsewhere<sup>48,49</sup>. Data matrices and R source code for replicating the central tests conducted on the data have been deposited online (<http://vm-lux.embl.de/~forslund/t2d/>).

**Evaluation of dietary habits.** A subset of the Danish study participants answered a validated food frequency questionnaire in order to obtain information on the habitual dietary habits. A complete data set was obtained for 66% of the nondiabetic individuals and 88% of T2D patients. When evaluating the dietary data, the consumed quantity was determined by multiplying portion size by the corresponding consumption frequency reported. Standard portion sizes for women and men, separately, were used in this calculation<sup>50,51</sup>. All food items in the questionnaire were linked to food items in the Danish Food Composition Databank<sup>52</sup>. Estimation of daily intake of macro- and micronutrients for each participant was based on calculations in the software program FoodCalc version 1.3<sup>53</sup>.

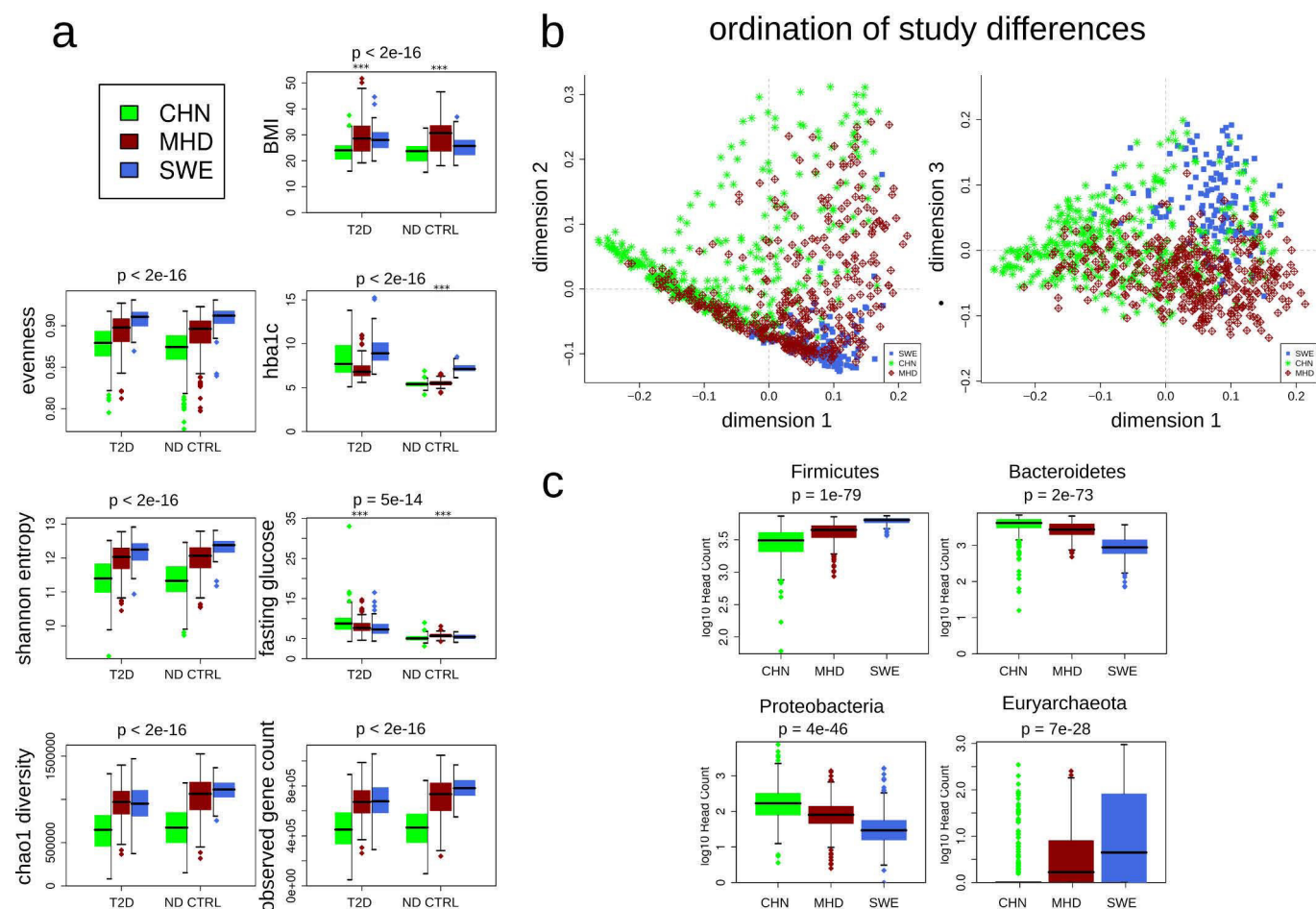
24. Jørgensen, T. *et al.* A randomized non-pharmacological intervention study for prevention of ischaemic heart disease: baseline results Inter99. *Eur. J. Cardiovasc. Prev. Rehabil.* **10**, 377–386 (2003).
25. WHO. Definition, *Diagnosis and Classification of Diabetes Mellitus and its Complications. Part 1: Diagnosis and Classification of Diabetes Mellitus*. Report No. WHO/NCD/NCS/99.2 (World Health Organization, 1999).
26. Li, J. *et al.* An integrated catalog of reference genes in the human gut microbiome. *Nature Biotechnol.* **32**, 834–841 (2014).
27. Kultima, J. R. *et al.* MOCAT: a metagenomics assembly and gene prediction toolkit. *PLoS ONE* **7**, e47656 (2012).
28. Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).
29. Arumugam, M., Harrington, E. D., Foerstner, K. U., Raes, J. & Bork, P. SmashCommunity: a metagenomic annotation and analysis tool. *Bioinformatics* **26**, 2977–2978 (2010).
30. Kanehisa, M. *et al.* KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* **36**, D480–D484 (2008).
31. Powell, S. *et al.* eggNOG v3.0: orthologous groups covering 1133 organisms at 41 different taxonomic ranges. *Nucleic Acids Res.* **40**, D284–D289 (2012).
32. Sunagawa, S. *et al.* Metagenomic species profiling using universal phylogenetic marker genes. *Nature Methods* **10**, 1196–1199 (2013).
33. Sunagawa, S. *et al.* Structure and function of the global ocean microbiome. *Science* **348**, (2015).
34. Nielsen, H. B. *et al.* Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nature Biotechnol.* **32**, 822–828 (2014).
35. Hildebrand, F. *et al.* LotuS: an efficient and user-friendly OTU processing pipeline. *Microbiome* **2**, 30 (2014).
36. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods* **10**, 996–998 (2013).
37. Edgar, R. C. *et al.* UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**, 2194–2200 (2011).
38. Magoč, T. & Salzberg, S. L. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957–2963 (2011).
39. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).
40. Madden, T. In *The NCBI Handbook [Internet]*. (eds, McEntyre J. & Ostell J.) Ch. 16 (National Center for Biotechnology Information, 2002) <http://www.ncbi.nlm.nih.gov/books/NBK21097/>
41. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. A Stat. Soc.* **57**, 289–300 (1995).
42. Hothorn, T., Hornik, K., van de Wiel, M. A. & Zeileis, A. A Lego system for conditional inference. *Am. Stat.* **60**, 257–263 (2006).
43. Dixon, P. VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* **14**, 927–930 (2003).
44. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer, 2009).
45. Anderson, M. J. A new method for non-parametric multivariate analysis of variance. *Austral. Ecol.* **26**, 32–46 (2001).
46. Friedman, J. *et al.* Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **33**, 1–22 (2010).
47. Abeel, T., Helleputte, T., Van de Peer, Y., Dupont, P. & Saey, Y. Robust biomarker identification for cancer diagnosis with ensemble feature selection methods. *Bioinformatics* **26**, 392–398 (2010).
48. Hildebrand, F. *et al.* A comparative analysis of the intestinal metagenomes present in guinea pigs (*Cavia porcellus*) and humans (*Homo sapiens*). *BMC Genomics* **13**, 514 (2012).
49. Hildebrand, F. *et al.* Inflammation-associated enterotypes, host genotype, cage and inter-individual effects drive gut microbiota variation in common laboratory mice. *Genome Biol.* **14**, R4 (2013).
50. Haraldsdóttir, J. *et al.* Portionsstorleker - Nordiska standardportioner av mat och livsmedel (Nordisk Ministerråd, 1998)
51. Biltoft-Jensen, A. *et al.* *Danskernes kostvaner 2000–2002*. DFVF publication No. 11 (Danmarks Fødevareforskning, Afdeling for Ernæring, 2005)
52. Møller, A. *et al.* Fødevaredatabanken version 5.0. Fødevareinformatik, Institut for Fødevaresikkerhed og Ernæring, Fødevaredirektoratet <http://www.foodcomp.dk> (2002).
53. Lauritsen, J. FoodCalc. [www.ibt.ku.dk/jesper/FoodCalc/](http://www.ibt.ku.dk/jesper/FoodCalc/) (2004).



**Extended Data Figure 1 | Validation of meta-analysis pipeline on simulated data.** **a**, As a positive control for the meta-analysis pipeline, true signal was removed from the data by randomly reshuffling sample labels. Artificial contrast was thereafter introduced between random groups containing as many such reshuffled samples as were in the original sets of T2D metformin+ ( $n_{\text{CHN}} = 15$ ,  $n_{\text{MHD}} = 58$ ,  $n_{\text{SWE}} = 20$ ) and T2D metformin- ( $n_{\text{CHN}} = 56$ ,  $n_{\text{MHD}} = 17$ ,  $n_{\text{SWE}} = 33$ ) samples in each original study subset, using the genus *Akkermansia* as an example feature. Samples randomly assigned to the sets of fake 'metformin-treated' and 'control' categories had their *Akkermansia* genus abundances adjusted to match the scale of the metformin effect on *Escherichia* genus abundance reported here (metformin-treated samples were roughly 150% as likely to have non-zero abundance, with a roughly threefold higher abundance where present), while retaining their data set origin labels. The full meta-analysis pipeline (study set blocked Kruskal–Wallis test, post-hoc Wilcoxon rank-sum test) was applied to these samples. Benjamini–Hochberg-corrected *P* values (FDR scores/*Q* values) from testing for a metformin effect on *Akkermansia* abundance are plotted in logarithmic scale on the vertical axis for 100 randomizations of the entire shuffled data set, either without (left box plot) or with (right box plot) the artificial *Akkermansia* metformin signal added after shuffling the data to remove original signal. Box plot borders show medians and quartiles, with points outside this range shown as vertical whisker lines and point markers. Whiskers extend to  $1.58 \times$  interquartile range/ $\sqrt{n}$ . Horizontal guide lines are shown for ease of visualization corresponding to different false discovery rate thresholds. For randomly reshuffled data, no significant contrast is detected as expected, whereas the artificially introduced signal is reliably detected,



roughly matching expectations from the definition of the false discovery rate itself. **b**, To investigate statistical power for the other medications tracked, five random sub-samplings were made of pairs of medicated and non-medicated samples at each increasing number of included sample pairs and the overall analysis was replicated for each. We tested each genus for significantly differential abundance between cases and controls (Kruskal–Wallis test followed by post-hoc Wilcoxon rank-sum test) at different Benjamini–Hochberg FDR significance cut-offs, which are represented by different colours. Of the total number of samples for which medication status was known, equal numbers (*n*) of medicated and unmedicated samples were chosen randomly in repeated iterations. This number *n* was varied up to its largest possible value (smallest of either number of medicated or unmedicated samples in the overall data set) and is shown on the x axis. The y axis shows the number of significant features relative to each cut-off. Error bars show  $\pm 1$  s.d. of each set of five randomized samples. **c**, The graphs show *Intestinibacter* and *Escherichia* median and quartile abundances as box plots, whiskers extend to  $1.58 \times$  interquartile range/ $\sqrt{n}$ , with samples that are extreme relative to the interquartile range shown as point markers, and with samples below detection threshold (DT) plotted at  $y = 0$ , in 21 additional T2D metformin+ and 9 additional T2D metformin- samples. Differences in abundance between sample categories are significant (Wilcoxon rank-sum test, Benjamini–Hochberg FDR < 0.1). All samples in which *Intestinibacter* was detected fall among the 9 out of 30 untreated rather than the 21 out of 30 metformin-treated samples, consistent with severe depletion under treatment; whereas *Escherichia* abundances increase under treatment, likewise consistent with observations from the main data set.



**Extended Data Figure 2 | Differences in physiological variables and microbiome characteristics between gut metagenome sample sets.**

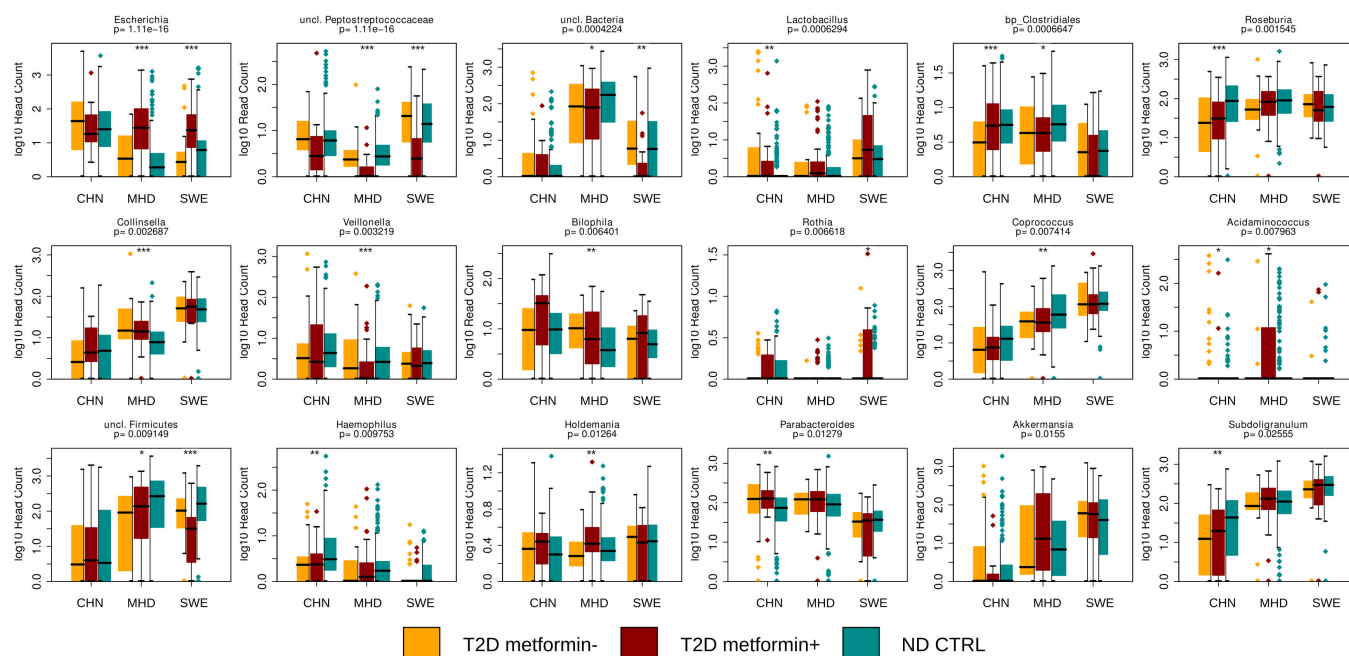
Chinese ( $n = 368$ ), Danish MetaHIT ( $n = 383$ ) and Swedish ( $n = 145$ ).

**a**, Several participant metadata variables are significantly different between cohorts. A subselection is shown as box plots displaying median and quartiles, with samples outside this range shown as point markers and whiskers. Whiskers extend to  $1.58 \times$  interquartile range/ $\sqrt{n}$ . **b**, In a principal coordinates analysis ordination of Bray–Curtis distances between samples on bacterial family level, clear differences between samples from the different cohorts become apparent. These are largely

explained by taxonomic differences as summarized at the phylum level.

**c**, Box plots for gut microbial taxa show medians and quartiles of log-transformed read counts for MOTUs summarized at the level of bacterial genera for the three country subsets across sample categories, with samples outside this range shown as point markers and whiskers. Whiskers extend to  $1.58 \times$  interquartile range/ $\sqrt{n}$ . For all box plots, tests for significant differences (Kruskal–Wallis test adjusted for study source) were performed, with  $P$  values shown at the head of each figure. Asterisks denote statistical significance of tests done for each country subset separately (\*\*\*) ( $P < 0.001$ ).





**Extended Data Figure 3 | Microbiome taxonomic composition comparison between gut metagenomes with particular focus on possible taxonomic restoration under metformin treatment for certain taxa.** T2D metformin- ( $n = 106$ ), T2D metformin+ ( $n = 93$ ) and ND control ( $n = 554$ ). Box plots show medians and quartiles log-transformed read counts for mOTUs summarized at the level of bacterial genera, for the

three country subsets across sample categories, with samples outside this range shown as point markers and whiskers. Whiskers extend to  $1.58 \times$  interquartile range/ $\sqrt{n}$ . Tests for significant differences (Kruskal–Wallis test adjusted for study source) were performed, with  $P$  values shown at the head of each figure. Asterisks denote statistical significance of tests for each country subset separately (\* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ).

Extended Data Table 1 | Analysis of variances

**a**

	Degrees of freedom	Sum of squares	Explained variation	F-statistic	Pr(>F)
Treatment	1	128	3.2%	21454.01	<2E-016 ***
Disease	1	42	1.1%	7039.86	<2E-016 ***
Country	1	376	9.4%	63206.96	<2E-016 ***
Treatment x Disease	1	1	0.0%	192.62	<2E-016 ***
Treatment x Country	1	67	1.7%	11209.33	<2E-016 ***
Disease x Country	1	1	0.0%	218.97	<2E-016 ***
Treatment x Disease x Country	1	0	0.0%	22.79	0.00000181 ***
Residuals	567001	3375	84.6%		
<b>Total:</b>		3990	100.0%		

**b**

Database OTU identifier	MWU P-value	Enriched in	Mean abundance (%)
OTU_45	0.048968332	T2D metformin+	0.803960725
OTU_1038	0.0319637913	T2D metformin-	0.000185722

Database OTU identifier	OTU_45	OTU_1038
Domain	<i>Bacteria</i>	<i>Bacteria</i>
Phylum	<i>Proteobacteria</i>	<i>Firmicutes</i>
Class	<i>Gammaproteobacteria</i>	<i>Clostridia</i>
Order	<i>Enterobacteriales</i>	<i>Clostridiales</i>
Family	<i>Enterobacteriaceae</i>	<i>Peptostreptococcaceae</i>
Genus	<i>Escherichia-Shigella</i>	<i>Intestinibacter</i>

**a.** The analysis of variances table shows the results of modelling the Canberra distances between T2D metformin- ( $n = 106$ ), T2D metformin+ ( $n = 93$ ) and ND control ( $n = 554$ ) samples with predictor variables encoding same/different diabetes status, same/different treatment, and same/different study source/country. Fractions of explained variance are taken as fractions of sum of square deviations from the model relative to the total deviation. **b.** Bacterial taxa found significantly different in gut abundance under metformin treatment were tested for significant differential relative abundance in a separate cohort under 16S amplicon sequencing between T2D metformin+ ( $n = 21$ ) and T2D metformin- ( $n = 9$ ) samples (Wilcoxon rank-sum test).

# Barcoding reveals complex clonal dynamics of *de novo* transformed human mammary cells

Long V. Nguyen<sup>1\*</sup>, Davide Pellacani<sup>1,2\*</sup>, Sylvain Lefort<sup>1</sup>, Nagarajan Kannan<sup>1,3</sup>, Tomo Osako<sup>3,4</sup>, Maisam Makarem<sup>1</sup>, Claire L. Cox<sup>1</sup>, William Kennedy<sup>1</sup>, Philip Beer<sup>1</sup>, Annaick Carles<sup>5</sup>, Michelle Moksa<sup>5</sup>, Misha Bilenky<sup>5,6</sup>, Sneha Balani<sup>1</sup>, Sonja Babovic<sup>1</sup>, Ivan Sun<sup>7,8</sup>, Miriam Rosin<sup>7,8</sup>, Samuel Aparicio<sup>3,4</sup>, Martin Hirst<sup>5,6</sup> & Connie J. Eaves<sup>1,2</sup>

**Most human breast cancers have diversified genomically and biologically by the time they become clinically evident<sup>1–3</sup>. Early events involved in their genesis and the cellular context in which these events occur have thus been difficult to characterize. Here we present the first formal evidence of the shared and independent ability of basal cells and luminal progenitors, isolated from normal human mammary tissue and transduced with a single oncogene (*KRAS*<sup>G12D</sup>), to produce serially transplantable, polyclonal, invasive ductal carcinomas within 8 weeks of being introduced either subrenally or subcutaneously into immunodeficient mice<sup>4</sup>. DNA barcoding<sup>5,6</sup> of the initial cells revealed a dramatic change in the numbers and sizes of clones generated from them within 2 weeks, and the first appearance of many ‘new’ clones in tumours passaged into secondary recipients. Both primary and secondary tumours were phenotypically heterogeneous and primary tumours were categorized transcriptionally as ‘normal-like’. This system challenges previous concepts that carcinogenesis in normal human epithelia is necessarily a slow process requiring the acquisition of multiple driver mutations. It also presents the first description of initial events that accompany the genesis and evolution of malignant human mammary cell populations, thereby contributing new understanding of the rapidity with which heterogeneity in their properties can develop.**

To investigate the susceptibility of different normal human mammary cell types to transformation under the influence of known oncogenes, we isolated CD49f<sup>+</sup>EpCAM<sup>low</sup> basal cells (BCs), CD49f<sup>+</sup>EpCAM<sup>+</sup> luminal progenitors (LPs), CD49f<sup>+</sup>EpCAM<sup>+</sup> non-clonogenic luminal cells (LCs), and non-epithelial CD49f<sup>+</sup>EpCAM<sup>+</sup> stromal cells (SCs) at high purity (>97%) by fluorescence-activated cell sorting (FACS) from 37 normal human reduction mammoplasty samples depleted of endothelial and haematopoietic cells<sup>4,7,8</sup> (Fig. 1a and Extended Data Tables 1 and 2). We then exposed one or more of these subsets to one or more oncogene-encoding lentiviral preparations (encoding complementary DNAs (cDNAs) for TP53<sup>R273C</sup> and green fluorescent protein (TP53<sup>R273C</sup>-GFP), PIK3CA<sup>H1047R</sup> and yellow fluorescent protein (PIK3CA<sup>H1047R</sup>-YFP)), and *KRAS*<sup>G12D</sup> and mCherry (*KRAS*<sup>G12D</sup>-mCherry) and, in some experiments, to a library of biologically neutral, barcoded lentiviral GFP vectors to allow subsequent clonal tracking of their progeny using a DNA sequencing approach<sup>5,6</sup> (Extended Data Fig. 1a). The cells were then embedded in a collagen gel ( $0.3 \times 10^5$  to  $16 \times 10^5$  cells per gel) and the gels transplanted into highly immunodeficient NOD-SCID *Il2rg*<sup>-/-</sup> (NSG) or NOD *Rag1*<sup>-/-</sup> *Il2rg*<sup>-/-</sup> (NRG) female mice.

In initial experiments,  $2 \times 10^5$  irradiated C3H-10T1/2 fibroblasts were co-embedded in the gels which were then transplanted subrenally, followed by subcutaneous implantation of the recipients with a slow-release capsule containing 17- $\beta$ -oestradiol and progestin (EP pellets)<sup>4</sup>, recognizing this would limit follow-up to 8 weeks because of incurred bladder toxicity. BCs and/or LPs isolated from 17 of 27 normal donors and exposed to all three oncogenic vectors produced tumours within 8 weeks (Fig. 1b) at similar overall frequencies (46% of BC isolates and 61% of LP isolates, respectively, Extended Data Fig. 1b). Identical treatment of LCs and SCs isolated from three of these samples did not produce any tumours in the same 8-week period. Both the BC- and LP-derived tumours resembled invasive ductal carcinomas (Fig. 2a and Extended Data Table 3) and were histologically very different from the organized bilayered structures generated in analogous xenografts of unmanipulated or simply barcoded normal human mammary cells<sup>4,6</sup>. Secondary female immunodeficient mice transplanted subcutaneously with a small piece (~25–33% of the initial tumour mass)<sup>5,9</sup> from four of nine of these primary tumours (one BC- and eight LP-derived) developed palpable tumours within another 8 weeks (Extended Data Table 4).

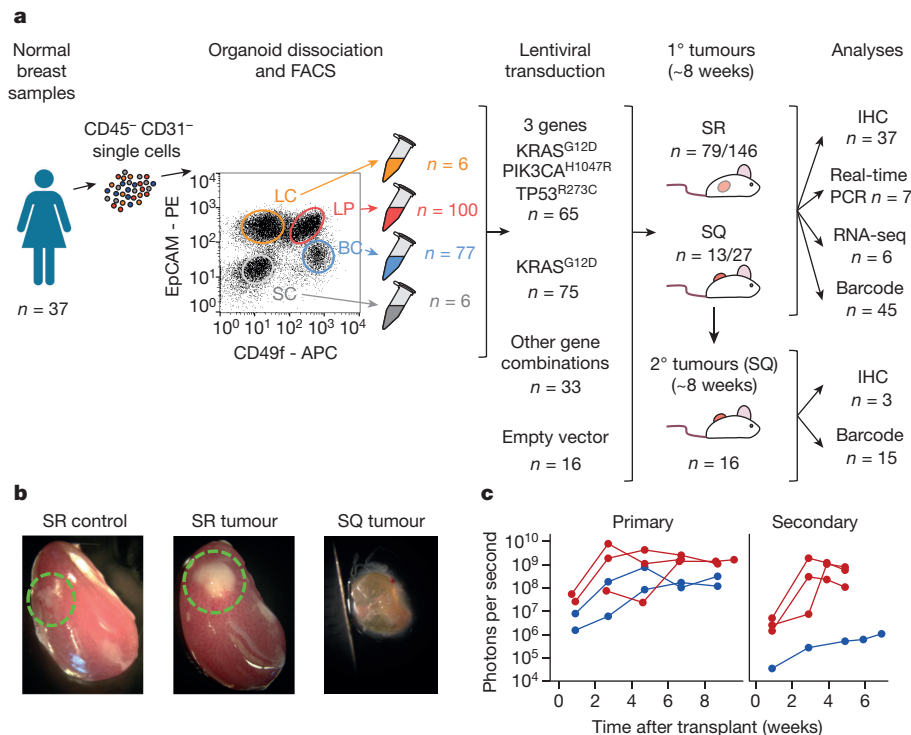
FACS analysis showed most of the transduced cells in all primary tumours examined co-expressed all three fluorescent reporters, consistent with the high transduction efficiencies measured in separate cell aliquots maintained *in vitro* for 72 h after virus exposure, and a similar expression profile was maintained in the single secondary tumour similarly analysed (Extended Data Fig. 1c, d). Notably, when the three oncogene-encoding viruses were tested on their own, or in pairs, tumours were obtained with similar efficiency only when the *KRAS*<sup>G12D</sup> vector was included, and even on its own (64/102 for all transductions that included *KRAS*<sup>G12D</sup> compared with 1/12 when *KRAS*<sup>G12D</sup> was not present; for more details, see Extended Data Fig. 1b and Extended Data Tables 1 and 2). PCR and Sanger sequencing confirmed the tumour cells contained the expected oncogene sequences including doubly and triply fluorescent cells isolated separately from tumours arising from cells initially exposed to three oncogenic vectors (Extended Data Fig. 2a–c).

Subsequent experiments demonstrated that invasive ductal carcinomas were also obtained at a similar frequency from both BCs and LPs (but not LCs or SCs from the same mammary samples) when the transduced cells were transplanted subcutaneously without irradiated fibroblasts or EP pellets, even when the cells were exposed only to the *KRAS*<sup>G12D</sup> vector (Extended Data Figs 1e and 2d and Extended Data Tables 1 and 2). These tumours could frequently also be serially passaged (Extended Data Fig. 2e and Extended Data Table 4) and their

<sup>1</sup>Terry Fox Laboratory, BC Cancer Agency, 675 West 10th Avenue, Vancouver, British Columbia V5Z 1L3, Canada. <sup>2</sup>Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia V6T 2B5, Canada. <sup>3</sup>Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, British Columbia V6T 2B5, Canada. <sup>4</sup>Department of Molecular Oncology, BC Cancer Agency, 675 West 10th Avenue, Vancouver, British Columbia V5Z 1L3, Canada. <sup>5</sup>Centre for High-Throughput Biology, Department of Microbiology & Immunology, University of British Columbia, 2125 East Mall, Vancouver, British Columbia V6T 1Z4, Canada. <sup>6</sup>Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency, 675 West 10th Avenue, Vancouver, British Columbia V5Z 1L3, Canada. <sup>7</sup>Biomedical Physiology and Kinesiology, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia V5A 1S6, Canada. <sup>8</sup>Cancer Control Unit, BC Cancer Agency, 675 West 10th Avenue, Vancouver, British Columbia V5Z 1L3, Canada.

\*These authors contributed equally to this work.





**Figure 1 | De novo formation of tumours from normal human mammary BCs and LPs.** **a**, Experimental design ( $n$  = number of experiments performed or tumours obtained). **b**, Examples of tumours compared with a control transplant which showed no evidence of a

growth more accurately monitored by luciferase bioluminescence (Fig. 1c and Supplementary Table 1).

FACS analysis of 15 tumours showed that  $48 \pm 5\%$  of the cells were human EpCAM<sup>+</sup> and/or HLA<sup>+</sup>, with similar results for BC- and LP-derived tumours (Extended Data Fig. 1c, f). Immunohistochemical (IHC) analyses of tumour sections (Fig. 2a, b) showed 88% and 55% of primary BC- and LP-derived tumours contained  $>5\%$  ER $\alpha$ <sup>+</sup> cells (median = 58% and 8% ER $\alpha$ <sup>+</sup> cells, respectively), but none contained  $>2\%$  PR<sup>+</sup> cells. HER2<sup>+</sup> cells were present at similar frequencies (in 88% and 52% of BC- and LP-derived tumours, respectively). Frequencies of Ki67<sup>+</sup> cells ranged from 2% to 30%, with only one secondary tumour containing as many as 70% Ki67<sup>+</sup> cells. In contrast, cells expressing EGFR, MUC1 and K8/18 were prevalent in almost all tumours examined. High K5 expression, normally exclusive to BCs, was prevalent (median = 90% K5<sup>+</sup> cells) in most LP-derived tumours, and less (median = 5% K5<sup>+</sup> cells) in BC-derived tumours. Expression of CD44, a marker associated with undifferentiated epithelial cells, was also less prevalent in BC- compared with LP-derived tumour cells (median = 2% and 50%, respectively). K14, another marker of normal human BCs, was also variably detected in both BC- and LP-derived tumours.

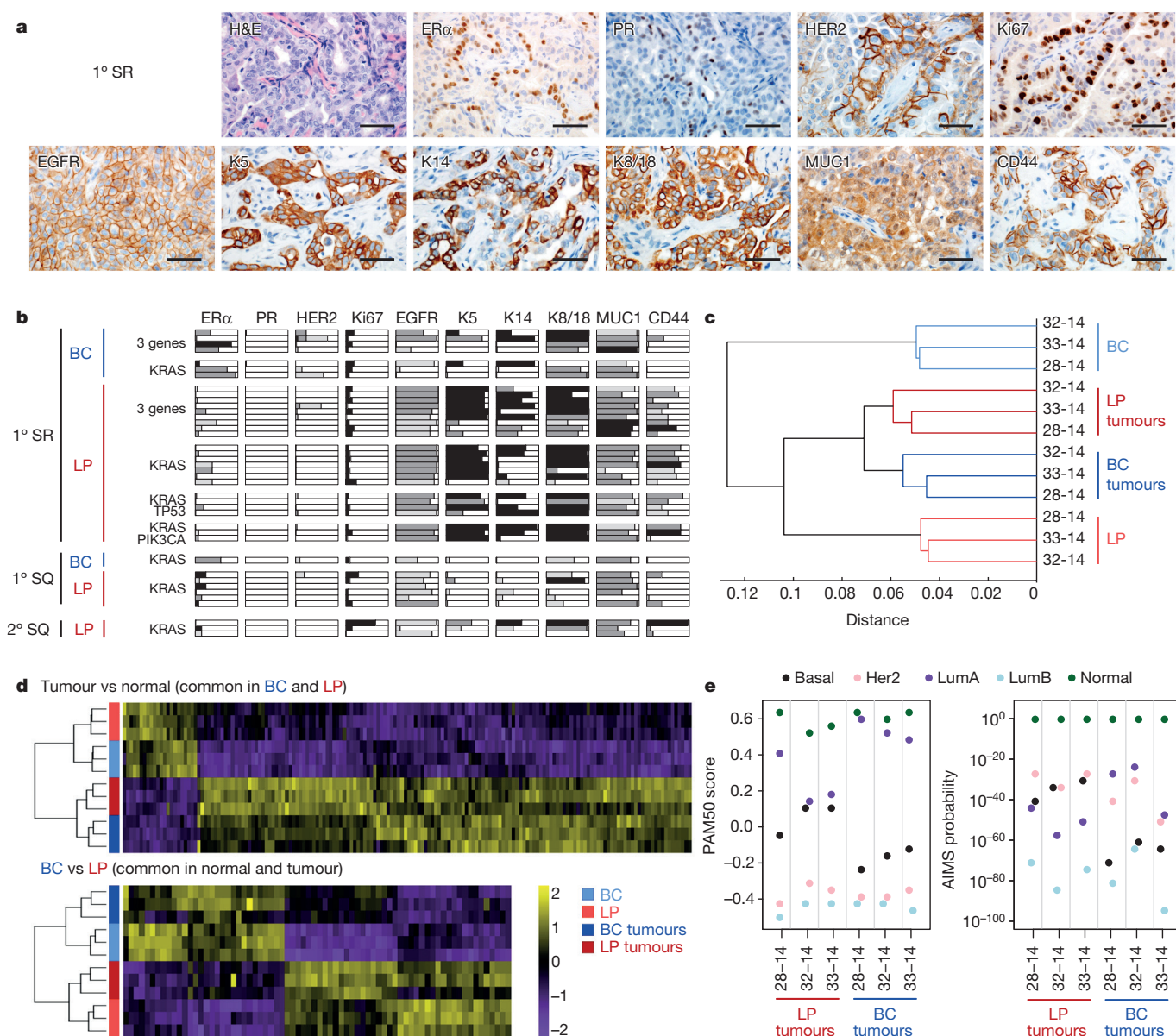
Gene expression analyses (Extended Data Fig. 2f) showed that transcripts for vimentin (*VIM*) and N-cadherin (*CDH2*), normally found exclusively in BCs, were present at high levels in both LP- and BC-derived tumours, with similar results for E-cadherin (*CDH1*) and *ELF5*, genes normally expressed exclusively in LPs and LCs. However, transcript levels of *SLUG* (*SNAI2*), another BC marker, were strongly decreased in BC-derived tumours whereas transcript levels of both *GATA3* and *NOTCH3*, two markers of LPs, showed little change. Cyclin-dependent kinase 1 (*CDK1*) was also highly expressed in all tumours, but other proliferation-associated genes, such as cyclin B1 (*CCNB1*) and *PCNA*, were highly expressed only in the LP-derived tumours. *TERT* transcripts remained at a similar level to that seen in the parental normal cell populations whereas those for *VEGFA*, *HIF1A* and *MAPK3* were more variable.

tumour when the mouse was killed. **c**, Bioluminescence measured over time in primary and secondary mice transplanted subcutaneously with oncogene- and luciferase-transduced BCs (blue) and LPs (red). SQ, subcutaneous; SR, subrenal.

RNA sequencing analysis was conducted on FACS-purified human cells isolated from six primary tumours generated from triply transduced cells (three from BCs and three from LPs) and the matched starting cells. Unsupervised clustering showed a closer relation of coding gene transcript levels in all six tumour populations to each other than to the normal cells from which the tumours had arisen (Fig. 2c). This prominent sharing of transcriptome changes in tumours derived both from BCs and from LPs suggests a key role of their mode of creation on their resultant molecular features. Specific differences in the gene expression changes that distinguished BC- and LP-derived tumours and their respective starting cell populations showed shared increased and decreased expression of 146 and 22 genes, respectively in both, indicative of a common gene signature in the transformants (Fig. 2d, top). Further analyses using either PAM50 (ref. 10) or AIMS classifier methodologies<sup>11</sup> indicated the transcriptional profiles of the *de novo* tumours most closely resembled those of spontaneous human breast cancers classified as 'normal-like' (Fig. 2e).

However, the unsupervised clustering also indicated that the three BC- and three LP-derived tumours formed separate groups, suggesting some retained influence of their different origins. This was further supported by the finding that  $>20\%$  (72) of the differentially expressed genes in the BC- and LP-derived tumours were similarly differentially expressed in the cells from which the tumours had arisen (Fig. 2d, bottom). Nevertheless, genes whose expression was upregulated in BC-derived tumours included several that are normally highly expressed in LPs and LCs but not BCs (for example, *AR*, *ESR1*, *FOXA1*, *TOX3*, *EPCAM*, *EHF* and *ELF5*). Conversely, the genes whose expression was upregulated in LP-derived tumours included several recognized BC-specific genes (for example, *VIM*, *TP63*, *ACTA2*, *THY1* and *CDH2*, Supplementary Table 2).

Clonal analyses were performed on primary tumours obtained from 45 isolates of BCs and LPs both from DNA extracted directly from tumour tissue and from FACS-purified human cells (Extended Data Fig. 3). The results showed a high variability in the clone content of different tumours (up to 1,700 using a threshold of 70 cells per



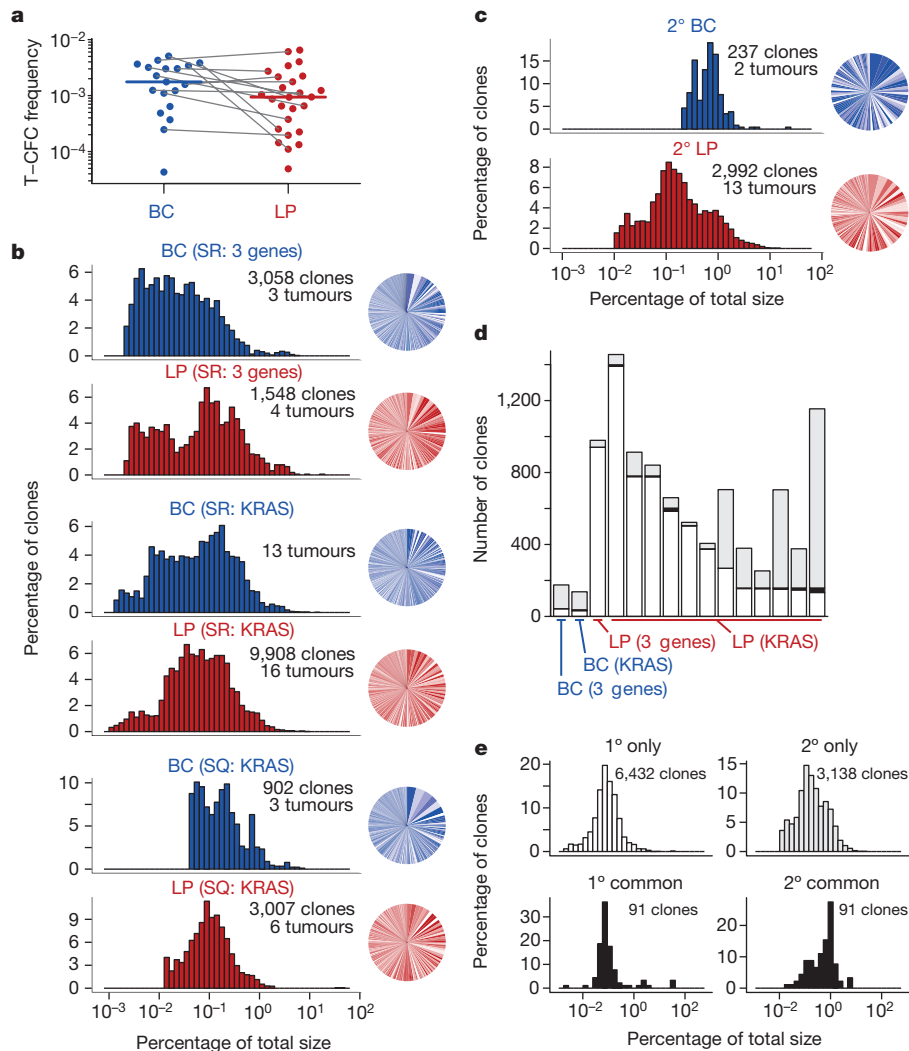
**Figure 2 | Phenotypic heterogeneity of primary and secondary *de novo* tumours.** **a**, Representative images of haematoxylin and eosin (H&E)- and antibody-stained sections from different primary BC- or LP- derived tumours arising from cells transplanted subrenally. Scale bar, 50  $\mu$ m. **b**, Prevalence of cells in each tumour analysed expressing the indicated marker at different levels: white, negative; light grey, dim; dark grey, mid; black, bright. Each bar represents a single tumour. **c**, Unsupervised

clustering (Spearman correlation and average linkage) of RNA sequencing data from *de novo* tumours compared with their matched cells of origin. **d**, Heat maps showing commonly differentially expressed genes between tumours and their matched cells of origin (top), and between BC- and LP-derived tumours in common with BCs and LPs (bottom). **e**, PAM50 and AIMS analyses of the RNA sequencing data from the *de novo* tumours.

clone), regardless of the protocol used for their generation (Extended Data Fig. 4a, b). Calculated (minimal) frequencies of tumorigenic clone-forming cells (T-CFCs) using the total number of initial cells transplanted as the denominator, ranged from 1/23,000 to 1/150. Paired comparisons for tumours produced from BCs and LPs from the same donor also did not reveal any effect on T-CFC frequency (Fig. 3a). To estimate clone sizes, we first derived 'relative' clone size values by normalizing each tumour to the sum of its absolute clone sizes. We then pooled the data for all tumours in each group being assessed. The overall distribution of relative clone sizes, like the clone frequencies, was very broad and showed no evidence of any effect of the cell of origin, oncogene(s) used or the transplantation site (Fig. 3b).

Analysis of 15 secondary tumours showed their clonal content was often high but again very variable, regardless of their origin (Fig. 3c and Extended Data Fig. 5a, b). Calculated frequencies of secondary

clones (with respect to the number of cells initially transplanted into primary mice) also yielded highly variable secondary T-CFC values but with no consistent difference from the calculated primary T-CFC frequencies. However, >75% of the clones detected in each secondary tumour were 'new'; that is, not detected in the matching primary tumour (Fig. 3d). Moreover, most of the clones present in multiple sibling secondary tumours produced from a common primary tumour were also different from one another (two primary tumours analysed, Extended Data Fig. 5c). Overall the total measured T-CFC frequencies (calculated from the total number of different clones in the primary or secondary tumours combined) ranged from  $\sim 1/5,700$  to  $\sim 1/120$  (Extended Data Fig. 5d). The relative sizes of the clones present in secondary tumours were also highly variable (Fig. 3c). Interestingly, in secondary tumours, the median size of the 'continuing' clones (evident in both primary and secondary tumours) was significantly larger



**Figure 3 | Barcoding reveals a complex clonal landscape of primary and secondary tumours.** **a**, Distribution of T-CFC frequencies (grey lines connect patient-matched BC- and LP-derived tumours, coloured lines show medians). **b**, Relative clone size distributions for all tumours in each group (left). Pie charts showing the clone size distributions in a representative tumour from each group (right). **c**, Distributions of the relative sizes of all clones pooled from all BC- and LP-derived

secondary tumours (left). Pie charts showing the clone size distributions in a representative tumour from each group (right). **d**, Clones detected exclusively in primary (white) or secondary (grey) tumours, or in both (black). **e**, Clone size distributions of combined primary or secondary tumours subdivided into those detected exclusively in either primary (white) or secondary (grey) tumours or both (black).

than the clones that first became detectable upon tumour passaging ( $P = 4 \times 10^{-12}$ , Mann–Whitney  $U$ -test, Fig. 3e right panels).

We then analysed the clonal composition of the cells produced from oncogene-transduced BCs and LPs after just 2 weeks in subrenal transplants, before tumours become grossly evident. The results showed the sizes as well as the numbers of clones detected at this time to be similar to those detected 6 weeks later in tumours derived from the same input cells (Fig. 4a and Extended Data Fig. 6a, b). The distributions of the relative clone sizes measured in the 2-week transplants both of BCs and of LPs were also similar (Fig. 4b). However, after 2 weeks, the absolute sizes of the clones derived from the  $KRAS^{G12D}$ -transduced LPs were already significantly larger than the sizes of the clones produced by matching transplants of control vector-transduced cells (median = 206 and 93, respectively,  $P = 3.3 \times 10^{-8}$ , Mann–Whitney  $U$ -test), with a slightly smaller effect apparent in the progeny of BCs from the same two donors (median = 112 and 94, respectively,  $P = 3.6 \times 10^{-7}$ , Mann–Whitney  $U$ -test, Fig. 4c).

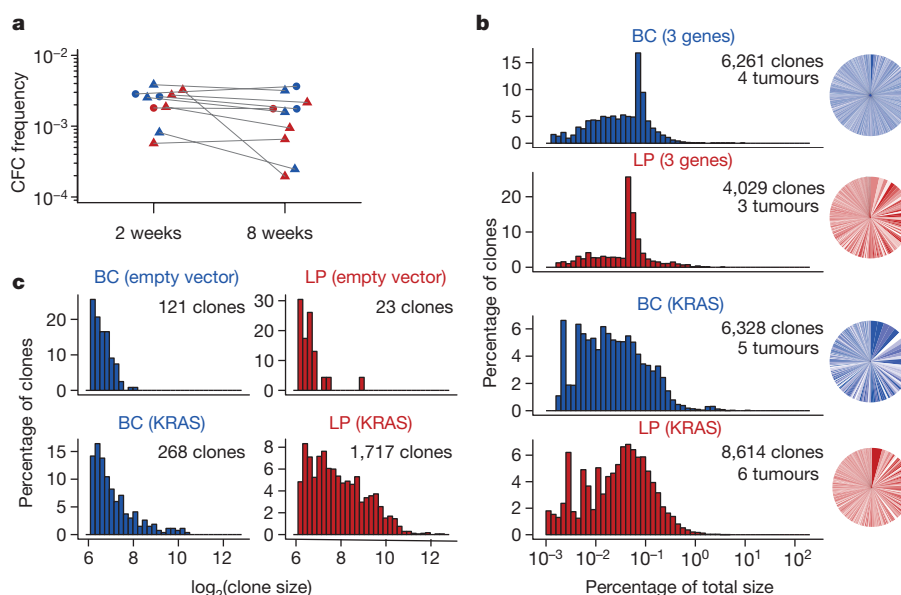
These studies provide new insights into the earliest phases of malignant transformation *in vivo* of cells isolated directly from normal human mammary tissue. Four findings are particularly noteworthy. The first is the rapidity and efficiency, albeit with high variability, with

which this process can be induced in prospectively purified, biologically distinct types of normal human mammary epithelial cells using a single transducing oncogene ( $KRAS^{G12D}$ ). This finding challenges previous assumptions of a requirement for a slow, multi-step selective process to accrue the genetic and/or epigenetic changes needed to obtain a continuously growing tumour. Interestingly, we did not obtain tumours from LCs or SCs subjected to the same protocols, in contrast to a recent report of highly  $ER\alpha^+$  tumours generated by transduction of  $EpCAM^+CD49f^-$  LCs with SV40/Ras<sup>12</sup>.

The second important finding was the considerable heterogeneity displayed in the numbers, phenotypes and growth behaviour of clonally tracked human cells with tumorigenic activity *in vivo* within 2–8 weeks. This result suggests that a similar range and speed of perturbations may accompany the spontaneous development of some breast cancers in patients.

A third and unexpected finding was the lack of a strong influence of the human mammary cell type initially transduced with the frequency of clones generated, the histopathology of the tumours produced or their loss of lineage-specific expression profiles. Taken together, this suggests a greater effect of the potent transforming role of the  $KRAS^{G12D}$  oncogene in these cells.





**Figure 4 | Early changes in clone growth in cells transduced with *KRAS*<sup>G12D</sup> only.** **a**, Comparison of the frequencies of BCs (blue) and LPs (red) that made clones in 2-week grafts (CFCs) versus those that made clones found in tumours 6 weeks later (T-CFCs). Cells transduced with all three oncogenes are shown as triangles, and with *KRAS*<sup>G12D</sup> only as circles (lines connect donor-matched samples). **b**, Distribution of relative clone

sizes for grafts containing BCs and LPs transduced with three oncogenes or *KRAS*<sup>G12D</sup> only (left). Pie charts showing the clone size distributions in a representative graft (right). **c**, Absolute clone size distributions of matched vector control-transduced BCs and LPs, and grafts containing BCs and LPs transduced with *KRAS*<sup>G12D</sup> ( $n = 2$  donors for each group).

The fourth finding was the frequent delayed activation of clonal growth observed in secondary tumours. This latency could either be biologically determined, reflecting an origin of these late appearing clones from their normal counterparts with similar features<sup>5</sup>, or simply reflective of a stochastic process, as previously indicated for established human breast cancer cell lines passaged *in vivo*<sup>4</sup>.

These results set the stage for examining the molecular basis of the biological heterogeneity now revealed that can occur during the earliest stages of breast cancer formation, the role of additional modifiers and how these may influence the acquisition of treatment response and resistance<sup>13,14</sup>.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 31 July 2014; accepted 22 September 2015.**

**Published online 2 December 2015.**

- Stephens, P. J. *et al.* Oslo Breast Cancer Consortium (OSBREAC). The landscape of cancer genes and mutational processes in breast cancer. *Nature* **486**, 400–404 (2012).
- Sørbye, T. *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl Acad. Sci. USA* **98**, 10869–10874 (2001).
- Curtis, C. *et al.* METABRIC Group. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346–352 (2012).
- Eirew, P. *et al.* A method for quantifying normal human mammary epithelial stem cells with *in vivo* regenerative ability. *Nature Med.* **14**, 1384–1389 (2008).
- Nguyen, L. V. *et al.* DNA barcoding reveals diverse growth kinetics of human breast tumour subclones in serially passaged xenografts. *Nature Commun.* **5**, 5871 (2014).
- Nguyen, L. V. *et al.* Clonal analysis via barcoding reveals diverse growth and differentiation of transplanted mouse and human mammary stem cells. *Cell Stem Cell* **14**, 253–263 (2014).
- Kannan, N. *et al.* Glutathione-dependent and -independent oxidative stress-control mechanisms distinguish normal human mammary epithelial cell subsets. *Proc. Natl Acad. Sci. USA* **111**, 7789–7794 (2014).
- Kannan, N. *et al.* The luminal progenitor compartment of the normal human mammary gland constitutes a unique site of telomere dysfunction. *Stem Cell Rep.* **1**, 28–37 (2013).

- Eirew, P. *et al.* Dynamics of genomic clones in breast cancer patient xenografts at single-cell resolution. *Nature* **518**, 422–426 (2015).
- Parker, J. S. *et al.* Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* **27**, 1160–1167 (2009).
- Paquet, E. R. & Hallett, M. T. Absolute assignment of breast cancer intrinsic molecular subtype. *J. Natl Cancer Inst.* **107**, 357 (2015).
- Keller, P. J. *et al.* Defining the cellular precursors to human breast cancer. *Proc. Natl Acad. Sci. USA* **109**, 2772–2777 (2012).
- Kreso, A. *et al.* Variable clonal repopulation dynamics influence chemotherapy response in colorectal cancer. *Science* **339**, 543–548 (2013).
- Bhang, H. E. *et al.* Studying clonal dynamics in response to cancer therapy using high-complexity barcoding. *Nature Med.* **21**, 440–448 (2015).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank D. Wilkinson, G. Edin and M. Hale for technical support, E. Bovill, J. Boyle, S. Bristol, P. Gdalevitch, A. Seal, J. Sproul and N. van Laeken for access to discarded reduction mammoplasty tissue, T. Nielsen and N. Poulin for discussions, the Centre for Translational and Applied Genomics (BC Cancer Agency) for assistance with IHC, and T. MacDonald for assistance with rodent husbandry. This work was supported by grants from the Canadian Cancer Society Research Institute, the Canadian Breast Cancer Foundation and the Canadian Breast Cancer Research Alliance. L.V.N. received a Vanier Canada Graduate Scholarship from the Canadian Institutes of Health Research (CIHR), and N.K. was supported by a MITACS Elevate Fellowship. T.O. was supported by a Molecular Oncologic Pathology Fellowship from CIHR and the Terry Fox Foundation, and by grants from the Sumitomo Life Welfare and Culture Foundation, the Mochida Memorial Foundation for Medical and Pharmaceutical Research, and the Takashi Tsuruo Memorial Fund. S.A. is supported by a Canada Research Chair.

**Author Contributions** L.V.N., D.P. and C.J.E. designed the project, drafted the manuscript and were assisted by S.L., C.L.C., W.K. and S. Balani in performing the experiments. M.M. and M.H. oversaw the generation of sequence data, and L.V.N., D.P., A.C. and M.B. analysed it. All authors contributed to the interpretation of the results, and read and approved the manuscript.

**Author Information** Final transcriptome data has been deposited in the European Genome-phenome Archive ([www.ebi.ac.uk/ega](http://www.ebi.ac.uk/ega)) under accession number EGAS00001001310. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.J.E. ([ceaves@bccrc.ca](mailto:ceaves@bccrc.ca)).

## METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

**Cells.** Reduction mammary tissue was collected with informed consent, as approved by the University of British Columbia Research Ethics Board, and dissociated to obtain organoid-rich pellets which were then viably cryopreserved<sup>4</sup>. Thawed organoids were rinsed with Hank's Balanced Salt Solution supplemented with 2% FBS (HF), and the cells then dissociated in 2.5 mg ml<sup>-1</sup> trypsin with 1 mM EDTA and 5 mg ml<sup>-1</sup> dispase (STEMCELL Technologies) with 100 µg ml<sup>-1</sup> DNaseI (Sigma), washing with HF between each step. The resulting cell suspension was filtered through a 40 µm mesh and BCs isolated by FACS according to their CD45<sup>-</sup>CD31<sup>-</sup>EpCAM<sup>hi</sup>CD49f<sup>+</sup> (or CD45<sup>-</sup>CD31<sup>-</sup>CD10<sup>+</sup>CD90<sup>+</sup>CD49f<sup>+</sup>) phenotype, LPs according to their CD45<sup>-</sup>CD31<sup>-</sup>EpCAM<sup>hi</sup>CD49f<sup>+</sup> (or CD45<sup>-</sup>CD31<sup>-</sup>CD10<sup>-</sup>CD90<sup>-</sup>CD49f<sup>+</sup>) phenotype, LCs according to their CD45<sup>-</sup>CD31<sup>-</sup>EpCAM<sup>hi</sup>CD49f<sup>-</sup> phenotype and SCs according to their CD45<sup>-</sup>CD31<sup>-</sup>EpCAM<sup>-</sup>CD49f<sup>-</sup> phenotype. In each case, a small aliquot of cells was immediately re-analysed to measure the purity of each sorted population (routinely >97%). Supplementary Table 3 lists the fluorochrome-labelled antibodies used.

**Lentiviral constructs and transduction.** Variations of the MNDU3-PGK-GFP lentiviral construct<sup>15</sup> encoding YFP or mCherry in place of the GFP reporter were generated and KRAS<sup>G12D</sup>, PIK3CA<sup>H1047R</sup> and TP53<sup>R273C</sup> mutant cDNAs then cloned into these, using flanking AscI and PacI restriction sites downstream of the MNDU3 promoter. Human KRAS cDNA was cloned from a human cell line, and altered by site-directed mutagenesis to obtain the G12D mutant. The TP53<sup>R273C</sup> mutant was cloned directly from a human cell line harbouring this mutation, and the human PIK3CA<sup>H1047R</sup> cDNA was obtained from A. Weng (Terry Fox Laboratory, BC Cancer Agency, Vancouver, BC, Canada). All cDNAs were sequence-verified before ligation into the lentiviral constructs. Clones confirmed to contain the mutant genes in the correct orientation were then selected for plasmid purification. Lentiviral supernatants containing ~10<sup>9</sup> infectious units per millilitre were produced<sup>16</sup> and added at a final dilution of 1:100 (for each) to cell suspensions containing 1 × 10<sup>6</sup> to 2 × 10<sup>6</sup> cells per 100 µl. The library of barcoded lentiviruses (titre of ~10<sup>9</sup> infectious units per millilitre)<sup>6</sup> was added to cells at a 1:200 dilution (~5 × 10<sup>5</sup> infectious units per 100 µl), to achieve an ~30% transduction efficiency.

**Xenografts.** Transduced human mammary epithelial cells were suspended in a neutralized rat tail collagen preparation<sup>4,17</sup> with 2 × 10<sup>5</sup> irradiated (15 Gy) C3H-10T1/2 mouse fibroblasts as indicated per 20 µl gel and the gels allowed to solidify at 37 °C for 30 min (refs 4, 17) before being implanted either subrenally or subcutaneously in 5- to 8-week-old virgin female NSG or NRG mice, that were then also implanted under the dorsal skin with a slow-release EP pellet, as indicated. Mice were bred, maintained and followed under specific-pathogen-free conditions in the Animal Resource Centre in the British Columbia Cancer Research Centre in accordance with protocols approved by the University of British Columbia Animal Care Committee. This included the monitoring of tumour growth for the periods indicated, or their removal earlier whenever a tumour reached a size of 1 cm<sup>3</sup>. The length along the tumour's longest axis was measured at the time of removal, and was immediately categorized as small (<5 mm), medium (5–10 mm) or large (10–15 mm). No other measurements were recorded. For some of the primary tumours, a small mechanically dissociated fragment or an enzymatically dissociated cell suspension was prepared and transplanted subcutaneously with 50% (v/v) matrigel, with or without an EP pellet, as indicated, into secondary recipients. To measure tumour bioluminescence from luciferase expression, mice were injected intraperitoneally with 150 mg kg<sup>-1</sup> body weight D-luciferin (Promega) and 10 min later imaged using a Xenogen IVIS Lumina system with Living Image version 3.0 software (Caliper Life Sciences). To prepare cell suspensions from tumours, the tissue was minced with a scalpel, incubated at 37 °C in DMEM/Ham's F12 media, supplemented with 5% FBS and 300 U ml<sup>-1</sup> collagenase and 100 U ml<sup>-1</sup> hyaluronidase for 2–4 h with periodic vortexing, washed with HF, and treated with 2.5 mg ml<sup>-1</sup> trypsin with 1 mM EDTA and 5 mg ml<sup>-1</sup> dispase with 100 µg ml<sup>-1</sup> DNaseI. Human cells were sorted after staining with anti-human-specific antibodies directed against EpCAM and HLA with simultaneous depletion of mouse cells stained with anti-mouse-specific antibodies directed against BP-1, CD140a, CD45 and CD31 (Supplementary Table 3).

**Histopathology and immunohistochemistry.** Collagen gels or pieces of tumours obtained from mice were fixed in 10% buffered formalin (Fisher), washed in 70% ethanol and embedded in paraffin. A tissue microarray using 1.5 mm dual cores per sample was constructed for all primary tumours analysed. From this tissue microarray, 4 µm sections were obtained. All secondary tumours were embedded in paraffin and sections prepared directly from these blocks. The sections of both primary and secondary tumours were either stained directly with H&E, or were first treated with Target Retrieval solution (DAKO) and then a cytation serum-free protein block (DAKO) followed by staining with either an anti-K14

antibody, an anti-MUC1 antibody, an anti-K5 antibody, an anti-K8/18 antibody, an anti-ER antibody, an anti-CD44 antibody, an anti-Ki67 antibody, an anti-HER2 antibody, an anti-PR antibody, or an anti-EGFR antibody. Use of a secondary rabbit antibody conjugated to horseradish peroxidase and treatment with 3,3'-diaminobenzidine (DAB, DAKO) was used to obtain a positive brown staining. Supplementary Table 3 provides details of the antibodies used and their sources. A negative control using one of the tumour samples, and a positive control using normal reduction mammary tissue, was included for each marker analysed. Tumour sections stained with H&E were reviewed for histomorphological analysis. Each of the tumours was classified into one of the histological subtypes according to the World Health Organization human breast tumour classification<sup>18</sup>. Nuclear grade was scored on a three-point scale according to the National Surgical Adjuvant Study of Breast Cancer (NSAS-BC) grading system<sup>19</sup>, and tubular formation, nuclear atypia, mitotic counts and histological grade were scored on a three-point scale according to the Nottingham grading system<sup>20</sup>. For each of the markers examined by IHC, intensity of staining (weak, intermediate or strong) and percentage of stained invasive tumour cells (0–100%) were scored. All of the histological and IHC parameters were scored blinded to the sample identity by a trained pathologist (T.O.).

**Barcode analysis.** Extracted genomic DNA was transferred to a 96-well plate in which researchers were blinded to their identity, although the order of the samples was not randomized. The samples were then treated identically, as previously described<sup>6</sup>. Defined numbers of control cells (10<sup>2</sup>–10<sup>6</sup>) containing a known barcode sequence at a single copy number per cell were analysed alone, and following their addition to each experimental sample, where they served as an internal normalization standard from which the relationship between fractional read value and cell numbers could be derived to calculate clone sizes (in absolute cell numbers). A threshold corresponding to a fractional read value equivalent of 70 cells was applied.

**Real-time PCR.** Total RNA was extracted from cryopreserved tumour samples using a mirVana miRNA isolation kit (Life Technologies) and cDNA then synthesized using SuperScript II Reverse Transcriptase (Life Technologies). Real-time PCR was performed using a SYBR Green master mix (Applied Biosystems) and samples were run in triplicate with the custom-designed primers listed in Supplementary Table 4. Human-specific primers were validated before use by testing for lack of reactivity with RNA from a panel of mouse tissues. Test gene cycle threshold (C<sub>t</sub>) values were normalized by subtracting the geometric mean of ΔC<sub>t</sub> values obtained for GAPDH and EIF4A1 (control) genes.

**RNA sequencing.** RNA was extracted from snap-frozen cells using the mirVana miRNA Isolation Kit (Life Technologies, AM1560) followed by ethanol precipitation. RNA was quantified using an Agilent Bioanalyzer (Life Technologies) and 100 ng of total RNA was ribosomal RNA (rRNA) depleted using a NEBNext rRNA Depletion Kit (New England BioLabs, E6310L). First strand cDNA was generated using a Maxima H minus First Strand cDNA Synthesis Kit (Thermo Scientific, K1652) with added Actinomycin D (1 µg, Sigma, A9415). The product was purified using in-house-prepared 20% PEG in 1 M NaCl Sera-Mag bead solution at a 1.8 × ratio and then eluted in 35 µl of Qiagen EB buffer. Second-strand cDNA was synthesized in a 50 µl volume using SuperScript Choice System for cDNA Synthesis (Life Technologies, 18090-019) with 12.5 mM GeneAmp dNTP Blend with dUTP. Double-stranded cDNA was purified with 20% PEG in 1 M NaCl Sera-Mag bead solution at a 1.8 × ratio and eluted in 40 µl of Qiagen EB buffer, and fragmented using Covaris E220 (55 s, 20% duty factor, 200 cycles per burst). Sheared cDNA was end repaired/phosphorylated, single A-tailed, and adaptor ligated using custom reagent formulations (New England BioLabs, E6000B-10) and in-house-prepared Illumina forked adaptor. PEG (20%) in 1 M NaCl Sera-Mag bead solution was used to purify the template between each of the enzymatic steps. To complete the process of generating strand directionality, adaptor-ligated template was digested with 5 U of AmpErase Uracil N-Glycosylase (Life Technologies, N8080096). Libraries were then indexed and PCR amplified using Phusion Hot Start II High Fidelity Polymerase (Thermo Scientific, F 549-L). An equal molar pool was sequenced on an Illumina MiSeq platform, which produced between 3 × 10<sup>6</sup> and 4 × 10<sup>6</sup> aligned sequence reads.

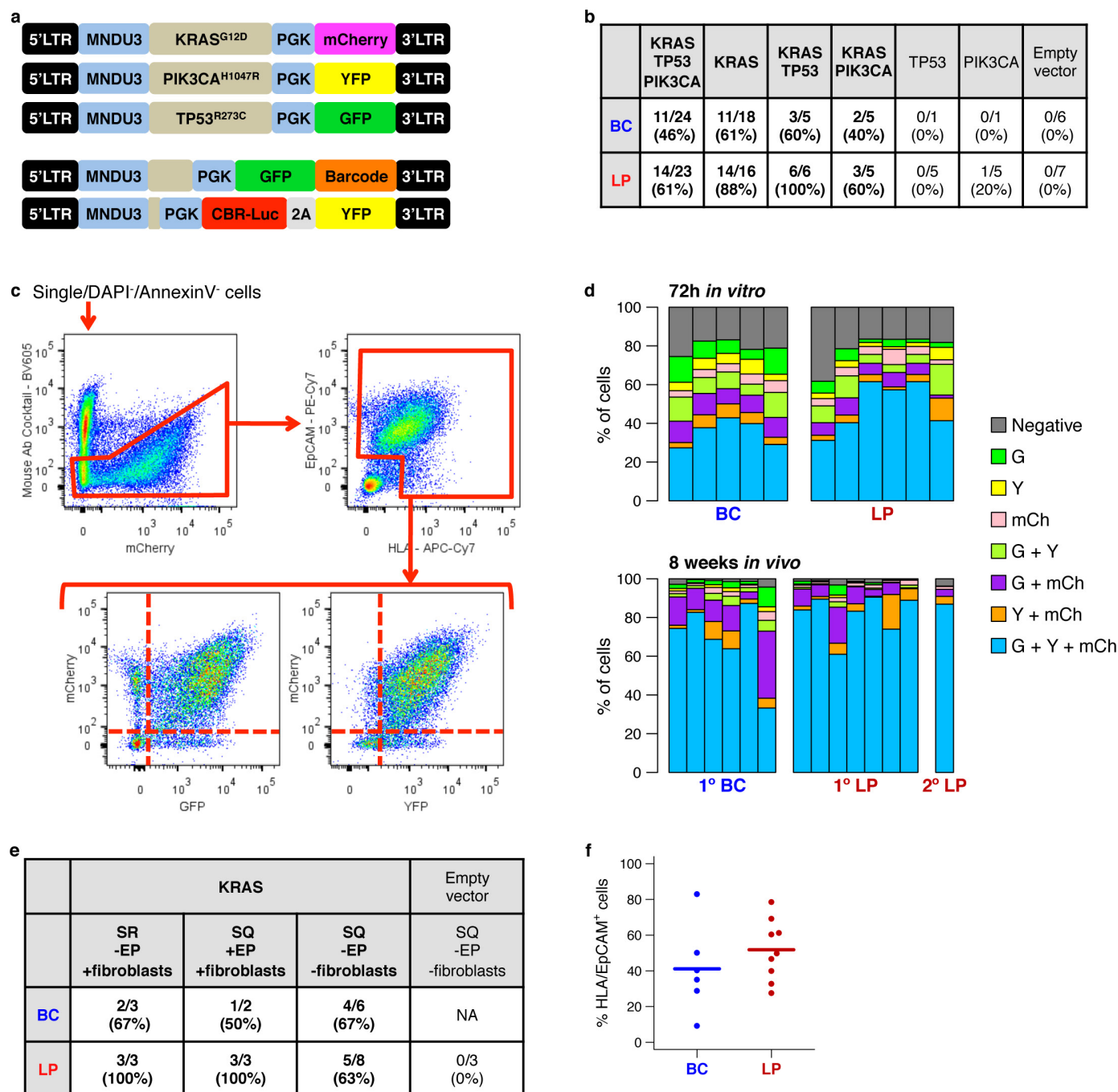
Adaptor sequences were stripped from the resulting 125-nucleotide sequence reads and the sequences uniformly trimmed to 75 nucleotides. Trimmed reads were aligned using BWA (version 0.5.7)<sup>21</sup> to a transcriptome reference<sup>22</sup> consisting of genomic sequence (GRCh37-lite July 2010) supplemented by read-length-specific exon–exon junction sequences. SAMtools (version 0.1.13)<sup>23</sup> was used to sort the alignment bam files. The sorted bam files were repositioned to GRCh37-lite using JAGuar (version 2.0.3)<sup>22</sup> to assign sequences that aligned across exon–exon junctions to their correct 'split' genomic coordinates. An in-house RNA quality control and analysis pipeline<sup>24</sup> was used to generate a report (Supplementary Table 5) and calculate a normalization constant for computing RPKM values (reads per kilobase per million mapped reads). The normalization constant was inferred from the total

number of exonic reads (excluding mitochondrial reads, reads from ribosomal genes and reads from highest 0.5% expressed exons).

Pairwise comparisons between different sample types within the same donor were performed to identify differentially expressed genes using a custom DEdefine matlab tool<sup>24</sup> (false discovery rate cutoff = 0.015, differentially expressed in at least two of three samples). PAM50 classification was performed following the methods described in ref. 10 using the R-Bioconductor script available at genome.unc.edu/pubsup/breastGEO/. AIMS classification<sup>11</sup> was performed using the Web tool available at [www.bci.mcgill.ca/AIMS/](http://www.bci.mcgill.ca/AIMS/).

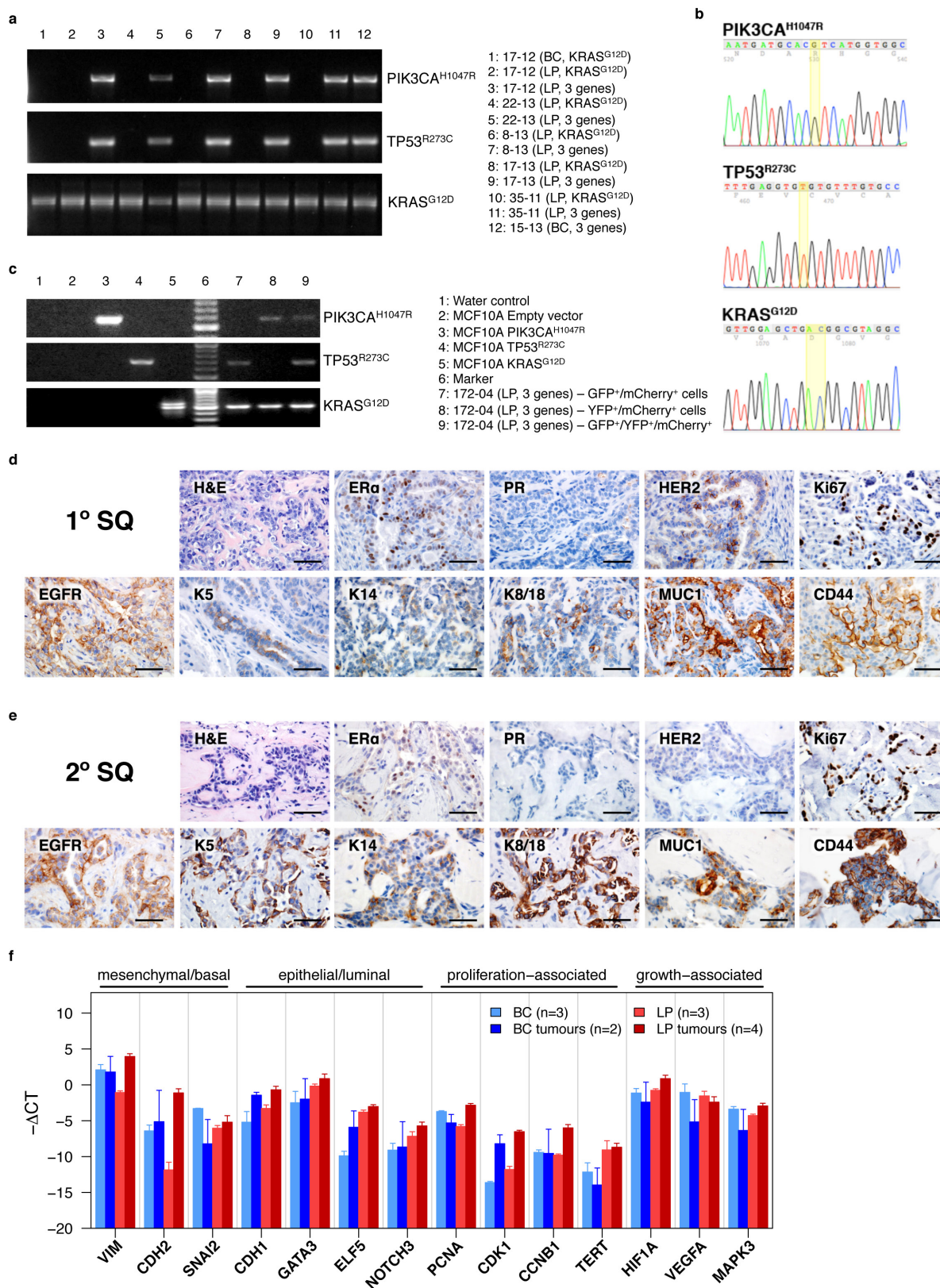
15. Logan, A. C. *et al.* Factors influencing the titer and infectivity of lentiviral vectors. *Hum. Gene Ther.* **15**, 976–988 (2004).
16. Imren, S. *et al.* High-level  $\beta$ -globin expression and preferred intragenic integration after lentiviral transduction of human cord blood stem cells. *J. Clin. Invest.* **114**, 953–962 (2004).
17. Eirew, P., Stingl, J. & Eaves, C. J. Quantitation of human mammary epithelial stem cells with *in vivo* regenerative properties using a subrenal capsule xenotransplantation assay. *Nature Protocols* **5**, 1945–1956 (2010).
18. Lakhani, S. R., Ellis, I. O., Schnitt, S. J., Tan, P. H. & van de Vijver, M. J. *WHO Classification of Tumours of the Breast* 4th edn, Ch. 2 (World Health Organization, 2012).
19. Tsuda, H., Akiyama, F., Kurosumi, M., Sakamoto, G. & Watanabe, T.; Japan National Surgical Adjuvant Study of Breast Cancer (NSAS-BC) Pathology Section. Establishment of histological criteria for high-risk node-negative breast carcinoma for a multi-institutional randomized clinical trial of adjuvant therapy. *Jpn. J. Clin. Oncol.* **28**, 486–491 (1998).
20. Elston, C. W. & Ellis, I. O. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* **19**, 403–410 (1991).
21. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
22. Butterfield, Y. S. *et al.* JAGuar: junction alignments to genome for RNA-seq reads. *PLoS One* **9**, e102398 (2014).
23. Li, H. *et al.* 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
24. Gascard, P. *et al.* Epigenetic and transcriptional determinants of the human breast. *Nature Commun.* **6**, 6351 (2015).





**Extended Data Figure 1 | Quantification of human cells containing different vector reporters in tumours derived from triply transduced starting populations. a**, Lentiviral constructs used. CBR-Luc, click beetle red luciferase. **b**, Frequencies of donor samples producing at least one tumour subrenally from BCs or LPs exposed to different combinations of oncogene-encoding vectors. **c**, Representative FACS profiles of a cell suspension prepared from a tumour produced from cells transduced with all three genes and sorted for human EpCAM and HLA using

human-specific antibodies. **d**, Percentage of cells expressing different lentiviral reporters in cells maintained *in vitro* for 72 h after transduction, and in primary and secondary tumours (G, GFP; Y, YFP; mCh, mCherry). **e**, Frequencies of donor samples producing at least one tumour under various transplantation conditions using BCs or LPs transduced with KRAS<sup>G12D</sup>. **f**, Percentages of human cells in BC- and LP-derived tumours detected by FACS on the basis of their expression of human EpCAM and/or HLA.



### Extended Data Figure 2 | Molecular characterization of the tumours.

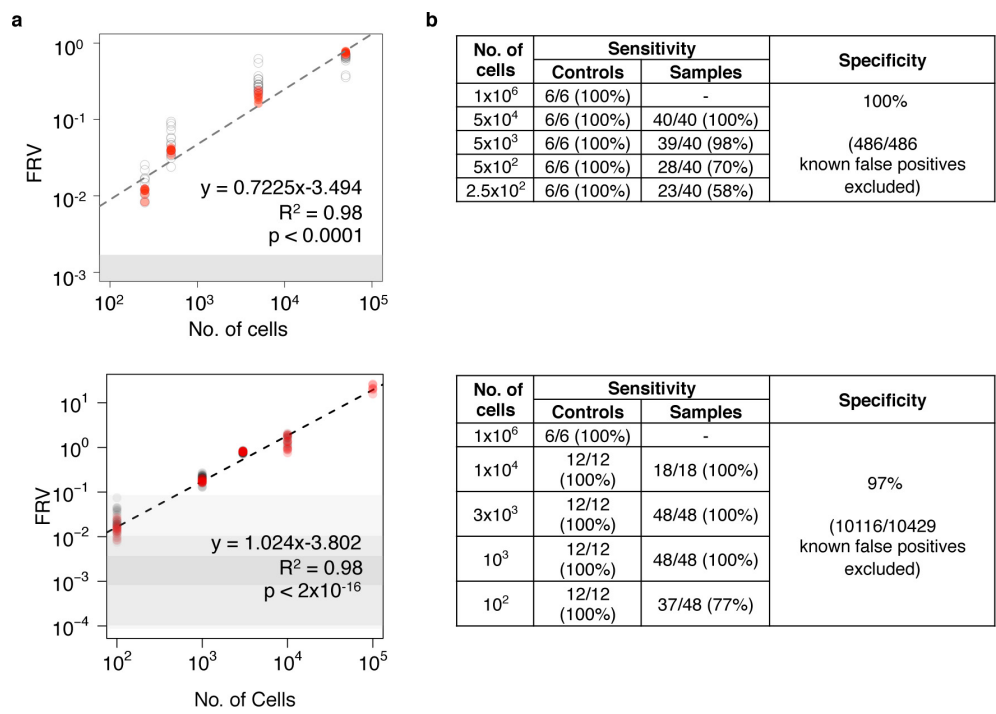
**a**, Examples of PCR evidence of all three vectors in DNA extracts obtained from a subset of tumours analysed with vector-specific primers.

**b**, A representative Sanger sequencing chromatograph showing the expected point mutations in the tumour cells analysed.

**c**, PCR evidence of the three vectors in FACS-purified doubly and triply transduced cells.

**d, e**, Representative images of H&E- and IHC-stained sections of

primary tumours (**d**, arising from cells transplanted subcutaneously) and secondary tumours (**e**, all arising subcutaneously) derived from either BCs or LPs. Scale bar, 50  $\mu$ m. **f**, Relative expression (negative  $\Delta C_t$  values, mean  $\pm$  s.e.m.) of gene transcripts typically associated with mesenchymal/basal or epithelial/luminal phenotypes, or associated with proliferation and cell growth.



**Extended Data Figure 3 | Threshold set for detection of barcoded clones for the two sequencing runs from which barcode data were acquired.**  
**a**, The relationship between the fractional read value (FRV) and the number of cells per clone. Spiked-in controls only and spiked-in controls

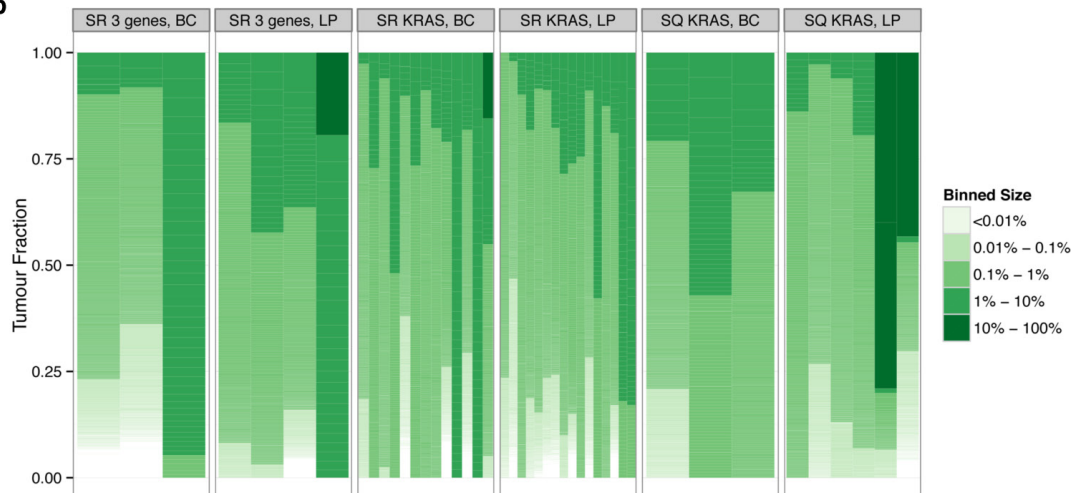
added to experimental samples are shown as red and grey points, respectively. The shaded grey box represents distribution of false positive barcodes. **b**, Sensitivity and specificity data for controls compared with experimental samples for different sized clones.



a

Donor	BC (3 genes)		BC (KRAS <sup>G12D</sup> )		LP (3 genes)		LP (KRAS <sup>G12D</sup> )	
	No. of clones	T-CFC freq.	No. of clones	T-CFC freq.	No. of clones	T-CFC freq.	No. of clones	T-CFC freq.
<b>SR +EP +Fibroblasts</b>								
14-13	1,335	1/570	1,200	1/630	941	1/570	1,155	1/460
18-13	1,680	1/270	1,462	1/310	-	-	1,108	1/1,100
15-13	43	1/23,000	-	-	-	-	-	-
17-12	-	-	267	1/900	224	1/4,500	376	1/2,700
17-13	-	-	-	-	355	1/1,500	645	1/960
25-11	-	-	-	-	28	1/20,000	332	1/1,700
31-13	-	-	1,650	1/440	-	-	883	1/1,500
21-13	-	-	171	1/230	-	-	1,399	1/160
27-14	-	-	149	1/800	-	-	602	1/900
33-14	-	-	612	1/260	-	-	277	1/4,000
33-14	-	-	37	1/4,100	-	-	84	1/5,100
33-14	-	-	311	1/2,100	-	-	-	-
28-14	-	-	447	1/1,600	-	-	-	-
23-14	-	-	37	1/2,700	-	-	-	-
8-13	-	-	-	-	-	-	959	1/290
22-13	-	-	-	-	-	-	862	1/810
13-12	-	-	-	-	-	-	375	1/720
15-14	-	-	-	-	-	-	65	1/7,50
<b>SR -EP +Fibroblasts</b>								
13-12	-	-	115	1/330	-	-	735	1/370
15-14	-	-	292	1/290	-	-	54	1/9,100
<b>SQ +EP +Fibroblasts</b>								
27-14	-	-	609	1/200	-	-	505	1/1,100
21-13	-	-	-	-	-	-	923	1/250
<b>SQ -EP -Fibroblasts</b>								
27-14	-	-	88	1/820	-	-	-	-
15-14	-	-	205	1/330	-	-	-	-
21-13	-	-	-	-	-	-	372	1/1,200
23-14	-	-	-	-	-	-	268	1/450
28-14	-	-	-	-	-	-	158	1/7,000
28-14	-	-	-	-	-	-	781	1/150

b

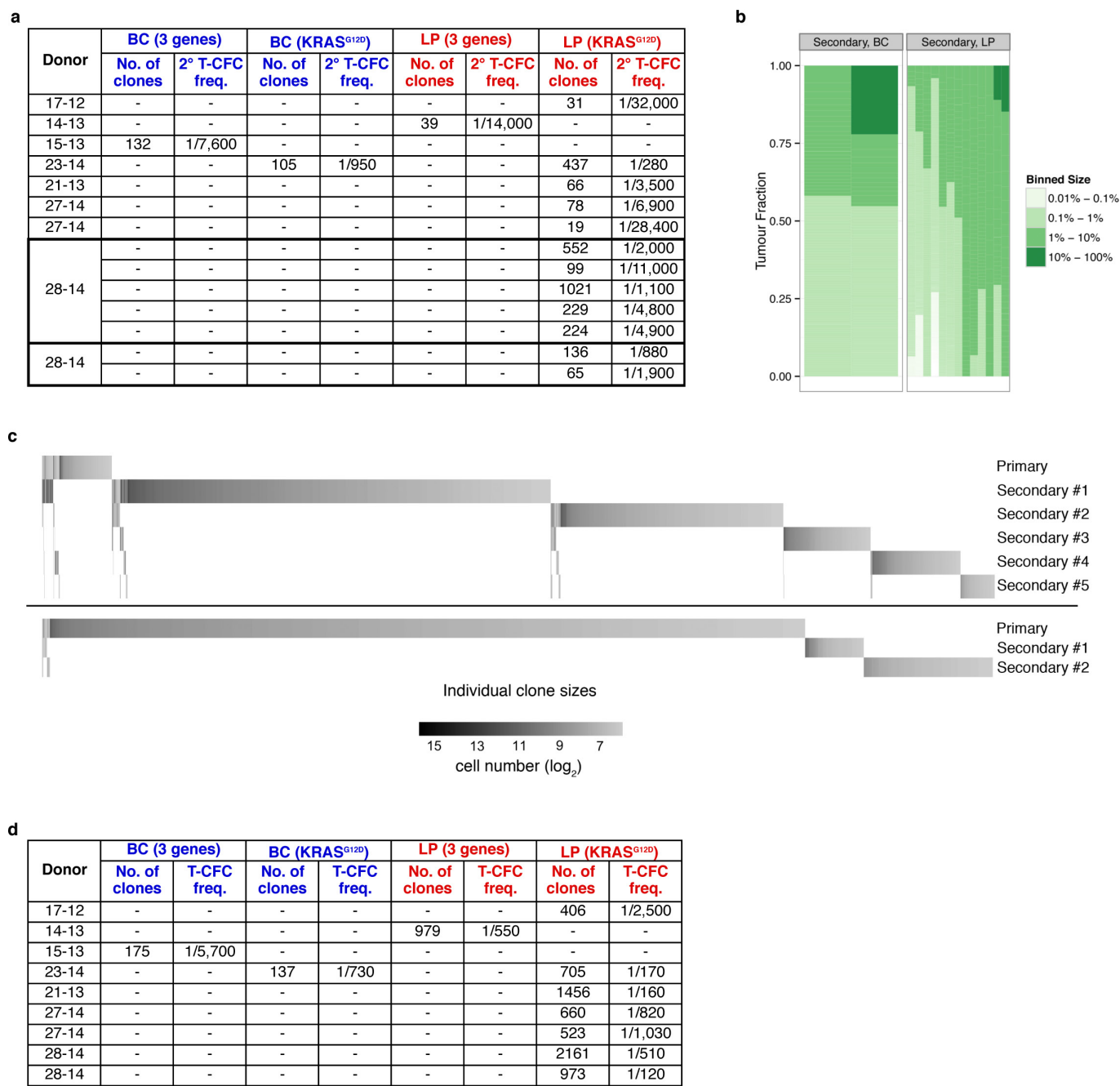


Extended Data Figure 4 | Clonal analyses of primary barcoded tumours.

a, Numbers of clones and frequencies of T-CFCs in primary tumours.

b, Relative clone size distributions for individual primary tumours grouped by the cell type initially manipulated and the oncogene(s) used.

Each column represents a single tumour. Each rectangle represents one clone. Its relative clone size is indicated by the shade of green, and its proportional contribution within each tumour is indicated by its length on the y axis.



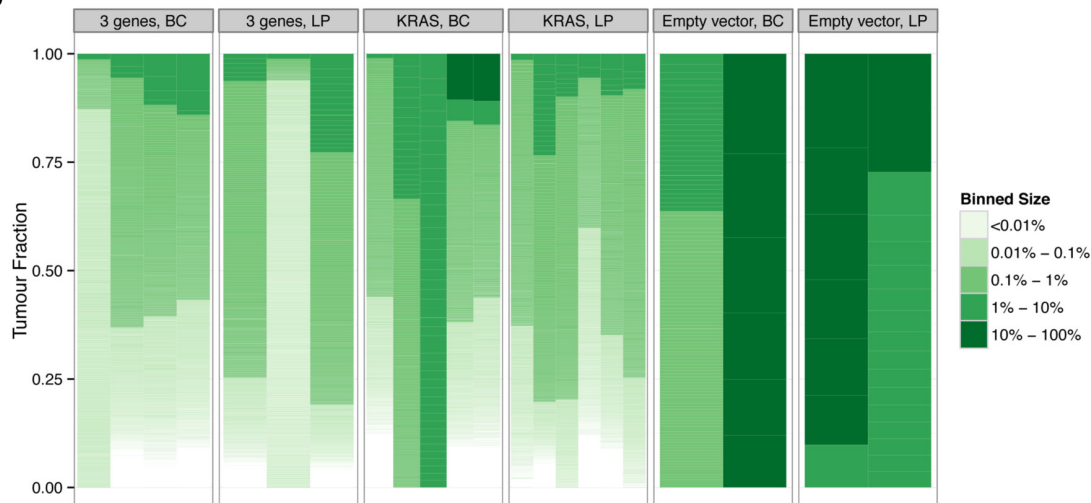
**Extended Data Figure 5 | Clonal analyses of secondary barcoded tumours.** **a**, Numbers of clones and frequencies of T-CFCs in secondary tumours. **b**, Relative clone size distributions for individual secondary tumours grouped by the cell type initially manipulated. Each column represents a single tumour. Each rectangle represents one clone. Its relative clone size is indicated by the shade of green, and its proportional contribution within each tumour is indicated by its length on the y axis.

**c**, Clonal landscape of replicate secondary tumours generated from single primary tumours in two separate experiments. Clones present in sibling tumours are shown above one another and unique clones are shown in the same horizontal bar. Increasing clone sizes are indicated by a grey intensity scale. **d**, Numbers of clones and T-CFC frequencies of combined primary and secondary tumours.

a

Donor	BC (3 genes)		BC (KRAS <sup>G12D</sup> )		BC (Empty vector)		LP (3 genes)		LP (KRAS <sup>G12D</sup> )		LP (Empty vector)	
	No. of clones	CFC freq.	No. of clones	CFC freq.	No. of clones	CFC freq.	No. of clones	CFC freq.	No. of clones	CFC freq.	No. of clones	CFC freq.
24-13	1,218	1/170	2,365	1/90			1,759	1/280	2,471	1/200		
14-13	1,988	1/380	1,929	1/390			966	1/550	1,467	1/360		
18-13	1,307	1/350	1,766	1/260			1,304	1/900	2,187	1/540		
31-13	1,748	1/420							772	1/1,700		
28-14			226	1/160	6	1/5,700			1,112	1/100	16	1/5,900
33-14			42	1/1,200	115	1/560			605	1/300	7	1/46,000

b



**Extended Data Figure 6 | Clonal analyses of transduced cells transplanted subrenally after 2 weeks *in vivo*.** a, Number of clones and frequency of CFCs in xenografts of transduced cells assessed after 2 weeks *in vivo*. b, Relative clone size distributions of individual 2-week transplants grouped by the cell type initially manipulated and the oncogene(s) used.

Each column represents a single transplant. Each rectangle represents one clone. Its relative clone size is indicated by the shade of green, and its proportional contribution within each tumour is indicated by its length on the y axis.



**Extended Data Table 1 | Primary xenotransplant experiments performed subrenally with EP pellets and irradiated fibroblasts**

3 Oncogenes, Primary SR, +EP +fibroblasts								
Patient ID	BC				LP			
	Cells transplanted (x10 <sup>3</sup> )	Weeks	Size	Analyses	Cells transplanted (x10 <sup>3</sup> )	Weeks	Size	Analyses
14-13	1000	7	S	B/C	1000	7	S	IHC, B/C, RT-PCR
172-04	40	8.5	M	IHC, RT-PCR	700	8.5	M	IHC, FACS
32-14	300	8	S	IHC	770	8	M	IHC, FACS, RNA-Seq
	770	8	S	FACS, RNA-Seq	300	8	S	FACS
28-14	260	8	S	IHC, FACS, RNA-Seq	740	8	M	FACS, RNA-Seq
23-14	70	8	S	H&E	500	8	M	IHC, FACS
33-14	360	8	S	FACS	470	8	M	IHC, FACS, RNA-Seq
	-	-	-	-	470	8	M	FACS
34-13	460	6	S	FACS	-	-	-	-
18-13	30	4	S	None	60	4	S	None
	450	8	S	B/C	-	-	-	-
15-13	1000	7	M	IHC, B/C, RT-PCR	-	-	-	-
22-13	100	8	S	None	-	-	-	-
31-13	730	8	S	None	1400	8	S	None
17-12	240	8	-	No tumour	1000	8	S	IHC, B/C, RT-PCR
17-13	120	8	-	No tumour	620	8	S	IHC, B/C, RT-PCR
25-11	120	8	-	No tumour	570	8	M	IHC, B/C, RT-PCR
64-06	-	-	-	-	358	8	S	None
	-	-	-	-	358	8	S	None
51-09	-	-	-	-	700	8.5	M	IHC
74-03	170	8	-	No tumour	50	8	S	H&E, FACS
	170	8	-	No tumour	503	8	-	No tumour
199-04	-	-	-	-	500	8.5	-	No tumour
55-07	1000	6	-	No tumour	1000	6	-	No tumour
36-04	1000	6	-	No tumour	1000	6	-	No tumour
52-08	400	8.5	-	No tumour	540	8.5	-	No tumour
143-10	890	8	-	No tumour	1200	8	-	No tumour
	-	-	-	-	1200	8	-	No tumour
28-13	960	8	-	No tumour	1600	8	-	No tumour
	-	-	-	-	1600	8	-	No tumour
129-10	110	8	-	No tumour	120	8	-	No tumour
33-13	285	8	-	No tumour	413	8	-	No tumour
	-	-	-	-	413	8	-	No tumour
32-13	220	8	-	No tumour	700	8	-	No tumour
	-	-	-	-	700	8	-	No tumour
	-	-	-	-	700	8	-	No tumour
8-13	130	8	-	No tumour	-	-	-	-

KRAS(G12D) only, Primary SR, +EP +fibroblasts								
17-12	240	8	S	B/C	1000	8	S	IHC, B/C
22-13	160	8	S	IHC	700	8	M	IHC, B/C
33-14	640	8	S	H&E, FACS, B/C	1100	8	S	FACS, B/C
	160	8	S	FACS, B/C	430	8	M	FACS, B/C
	150	8	S	FACS, B/C	-	-	-	-
27-14	120	8	S	IHC, FACS, B/C	540	8	S	IHC, FACS, B/C
21-13	40	8	S	H&E, FACS, B/C	230	8	M	H&E, FACS, B/C
14-13	760	8	S	B/C	540	8	S	B/C
31-13	730	8	S	B/C	1400	8	S	B/C
34-13	150	6	S	None	290	6	S	None
28-14	700	8	S	IHC, FACS, B/C	-	-	-	-
23-14	100	8	S	FACS, B/C	-	-	-	-
18-13	30	4	-	No tumour	60	8	S	None
	450	8	S	B/C	1000	8	M	IHC, B/C
17-13	120	8	-	No tumour	620	8	M	IHC, B/C
25-11	120	8	-	No tumour	570	8	S	IHC, B/C
13-12	38	8	-	No tumour	270	8	S	FACS, B/C
15-14	85	8	-	No tumour	490	8	S	FACS, B/C
8-13	-	-	-	-	280	8	M	IHC, B/C
29-14	65	8	-	No tumour	580	8	-	No tumour
74-06	200	8	-	No tumour	400	8	-	No tumour
	-	-	-	-	400	8	-	No tumour
	-	-	-	-	400	8	-	No tumour
	-	-	-	-	400	8	-	No tumour
33-13	63	9	-	No tumour	-	-	-	-

This table provides information for each experiment where transplants were performed subrenally with the addition of an EP pellet implant, and irradiated mouse fibroblasts were combined with the transduced human mammary cells in the transplanted collagen gel. Shown are the unique numbers assigned to donor-specific mammary samples, the subtypes of cells transduced and the oncogene(s) used for the transduction. Also shown are the sizes of the tumours obtained, and the subsequent analyses performed, where applicable. Small (S), medium (M) and large (L) refer to tumours for which the longest axis was measured as <5 mm, from 5 to 10 mm, or 10 to a maximum of 15 mm, respectively, to conform to our institutionally approved animal use protocols that stipulate tumours must be removed when they reach a size of 1 cm<sup>3</sup>. When an experiment was not performed or a data field was not applicable, this is indicated with '-'. When a transplant was performed and no tumour was produced, this is indicated as 'no tumour'. The subsequent analyses performed are indicated, where applicable (B/C, barcode analysis).

Extended Data Table 2 | Primary xenotransplant experiments testing different variables

KRAS(G12D) only, Primary SR, no EP, +fibroblasts (continued)								
Patient ID	BC				LP			
	Cells transplanted ( $\times 10^3$ )	Weeks	Size	Analyses	Cells transplanted ( $\times 10^3$ )	Weeks	Size	Analyses
13-12	38	16	S	FACS, B/C	270	9	M	FACS, B/C
15-14	85	16	S	FACS, B/C	490	9	S	FACS, B/C
33-13	63	9	-	No tumour	940	9	M	H&E
KRAS(G12D) only, Primary SQ, +EP +fibroblasts								
27-14	120	8	S	IHC, FACS, B/C	540	8	M	IHC, FACS, B/C
33-14	156	8	-	No tumour	1065	8	S	IHC
21-13	-	-	-	-	230	8	S	H&E, FACS, B/C
KRAS(G12D) only, Primary SQ, no EP, no fibroblasts								
38-14	10	7	S	FACS	135	9	-	No tumour
15-14	68	9	S	FACS, B/C	350	13	-	No tumour
27-14	72	9	M	H&E, FACS, B/C	405	13	-	No tumour
33-14	135	9	-	No tumour	210	7	S	FACS
	-	-	-	-	930	10	-	No tumour
42-14	280	9	-	No tumour	40	7	M	FACS
21-13	40	11	S	FACS	430	9	M	IHC, FACS, B/C
28-14	-	-	-	-	1100	9	M	IHC, FACS, B/C
	-	-	-	-	120	9	M	IHC, FACS, B/C
23-14	-	-	-	-	650	9	S	IHC, FACS, B/C
KRAS(G12D)+TP53(R273C), Primary SR, +EP +fibroblasts								
22-12	160	8	S	None	700	8	S	None
17-13	120	8	S	None	620	8	M	None
25-11	120	8	M	None	570	8	M	None
17-12	240	8	-	No tumour	1000	8	S	None
18-13	30	4	-	No tumour	60	4	S	None
8-13	-	-	-	-	280	8	L	None
KRAS(G12D)+PIK3CA(H1047R), Primary SR, +EP +fibroblasts								
17-12	-	-	-	-	1000	8	M	None
17-13	120	8	S	None	620	8	M	None
4-13	30	4	S	None	60	4	-	No tumour
22-13	160	8	-	No tumour	-	-	-	-
25-11	120	8	-	No tumour	570	8	M	None
8-13	130	8	-	No tumour	280	8	-	No tumour
TP53(R273C) only, Primary SR, +EP +fibroblasts								
17-12	240	8	-	No tumour	1000	8	-	No tumour
22-13	-	-	-	-	70	8	-	No tumour
17-13	-	-	-	-	620	8	-	No tumour
25-11	-	-	-	-	570	8	-	No tumour
8-13	-	-	-	-	280	8	-	No tumour
PIK3CA(H1047R) only, Primary SR, +EP +fibroblasts								
8-13	-	-	-	-	280	8	S	None
17-12	240	8	-	No tumour	1000	8	-	No tumour
22-13	-	-	-	-	70	8	-	No tumour
17-13	-	-	-	-	620	8	-	No tumour
25-11	-	-	-	-	570	8	-	No tumour

This table provides information for all other primary xenotransplant experiments not included in Extended Data Table 1. Shown are the unique numbers assigned to donor-specific mammary samples, the subtypes of cells transduced, the oncogene(s) used for the transduction, whether irradiated mouse fibroblasts were combined with transduced human mammary cells when a collagen gel was transplanted, the site of the transplant (subrenally or subcutaneously) and whether an EP pellet implant was used. Also indicated are the sizes of the tumours obtained, and the subsequent analyses performed, where applicable. Small (S), medium (M), and large (L) refer to tumours for which the longest axis was measured as <5 mm, from 5 to 10 mm, or 10 to a maximum of 15 mm, respectively, to conform to our institutionally approved animal use protocols that stipulate tumours must be removed when they reach a size of 1 cm<sup>3</sup>. When an experiment was not performed or a data field was not applicable, this is indicated with '-'. When a transplant was performed and no tumour was produced, this is indicated as 'no tumour'. The subsequent analyses performed are indicated, where applicable (B/C, barcode analysis).

Extended Data Table 3 | Histopathological characterization of the *de novo* tumours

3 Oncogenes, Primary SR, +EP +fibroblasts									
Patient ID	Starting cells	Histology	Fibrous Stroma	Nuclear grade	Histological grade	Tubular formation	Nuclear atypia	Mitotic index	Analyses
15-13	BC	Invasive ductal	1	2	1	1	2	2	H&E, IHC
172-04	BC	Invasive ductal	-	-	-	-	-	-	H&E, IHC
	LP	Invasive ductal	2	1	1	2	2	1	H&E, IHC
28-14	BC	Invasive ductal	2	1	1	2	2	1	H&E, IHC
32-14	BC	Invasive ductal	2	2	1	1	2	2	H&E, IHC
	BC	Invasive ductal	2	1	1	1	1	1	H&E, IHC
	LP	Invasive ductal	2	1	1	2	1	1	H&E
	LP	Invasive ductal	2	1	1	1	2	1	H&E
23-14	BC	Invasive ductal	2	1	1	1	1	1	H&E
	LP	Invasive ductal	1	1	1	2	2	1	H&E, IHC
14-13	LP	Invasive ductal	3	2	1	2	2	1	H&E, IHC
17-12	LP	Invasive ductal	3	1	1	1	2	1	H&E, IHC
17-13	LP	Invasive ductal	2	1	1	2	2	1	H&E, IHC
33-14	LP	Invasive ductal	2	1	1	2	2	1	H&E, IHC
25-11	LP	Invasive ductal	2	1	1	1	2	1	H&E, IHC
51-09	LP	Invasive ductal	2	1	1	2	2	1	H&E, IHC
74-03	LP	Invasive ductal	-	1	1	1	1	1	H&E
KRASG12D only, Primary SR, +EP +fibroblasts									
21-13	BC	Invasive ductal	2	1	1	2	1	1	H&E
	LP	Invasive ductal	2	1	1	2	1	1	H&E
22-13	BC	Invasive ductal	2	1	1	2	2	1	H&E, IHC
	LP	Invasive ductal	3	1	1	1	1	1	H&E, IHC
27-14	BC	Invasive ductal	1	1	1	2	2	1	H&E, IHC
	LP	Invasive ductal	2	1	1	2	1	2	H&E, IHC
28-14	BC	Invasive ductal	2	1	1	1	2	1	H&E, IHC
33-14	BC	Invasive ductal	3	1	1	1	1	1	H&E
17-12	LP	Invasive ductal	2	1	1	1	2	1	H&E, IHC
17-13	LP	Invasive ductal	3	1	1	2	2	1	H&E, IHC
18-13	LP	Invasive ductal	3	1	2	3	2	1	H&E, IHC
35-11	LP	Invasive ductal	3	1	1	1	1	1	H&E, IHC
8-13	LP	Invasive ductal	2	1	1	1	1	1	H&E, IHC
KRASG12D only, Primary SR, no EP +fibroblasts									
33-13	LP	Invasive ductal	2	2	2	2	2	2	H&E
KRASG12D only, Primary SQ, +EP +fibroblasts									
27-14	BC	Invasive ductal	2	1	1	2	2	1	H&E, IHC
21-13	LP	Invasive ductal	2	1	1	2	1	1	H&E
27-14	LP	Invasive ductal	2	1	1	2	1	1	H&E, IHC
33-14	LP	Invasive ductal	1	2	2	2	2	2	H&E, IHC
KRASG12D only, Primary SQ, no EP, no fibroblasts									
27-14	BC	Invasive ductal	1	1	2	3	1	1	H&E
21-13	LP	Invasive ductal	1	1	1	2	1	1	H&E, IHC
23-14	LP	Invasive ductal	3	2	2	2	2	2	H&E, IHC
28-14	LP	Invasive ductal	3	3	2	2	2	3	H&E, IHC
	LP	Invasive ductal	3	2	2	2	2	2	H&E, IHC
Secondary tumors (all SQ)									
21-13	LP	Invasive ductal	2	1	1	2	2	1	H&E, IHC
28-14	LP	Invasive ductal	3	2	2	2	2	2	H&E, IHC
	LP	Invasive ductal	3	1	1	2	2	1	H&E, IHC
	LP	Invasive ductal	-	1	1	1	1	1	H&E
	LP	Invasive ductal	3	2	1	1	2	2	H&E

This table summarizes the results on all tumours analysed according to the cell type initially transduced, the oncogenes used, the site of the transplant and whether EP or irradiated fibroblasts were used. The amount of fibrous stroma present was scored as low (1), intermediate (2) or high (3). Nuclear and histological features were used to score tumours as low-grade or well-differentiated (1), intermediate-grade or moderately differentiated (2) or high-grade or poorly differentiated (3). Tubular formation was scored as >75% (1), 10–75% (2), or <10% (3) of the tumour area forming glandular/tubular structures. Nuclear atypia was scored as increasing from 1 to 3 on the basis of evidence of slight to marked variation compared with normal tissue. Increasing mitotic counts were scored as 0–4 (1), 5–10 (2), or  $\geq 11$  (3) mitotic figures per 400 $\times$  high-power field. Tumours were also examined for evidence of necrosis, but none was found.



Extended Data Table 4 | Details of all secondary xenotransplant experiments

**3 Oncogenes, Secondary tumours (all SQ)**

Patient ID	Basal cells			Luminal Progenitor cells		
	Weeks	Size	Analyses	Weeks	Size	Analyses
15-13	8	M	B/C	-	-	-
14-13	-	-	-	8	M	B/C
18-13	-	-	-	8	M	None
64-06	-	-	-	7	S	FACS
17-13	-	-	-	8	-	No tumour
25-11	-	-	-	8	-	No tumour
23-14	-	-	-	9	-	No tumour
28-14	-	-	-	9	-	No tumour
33-14	-	-	-	9	-	No tumour
	-	-	-	9	-	No tumour

**KRAS(G12D) only, Secondary tumours (all SQ)**

23-14	7	S	FACS, B/C	5	S	FACS, B/C
27-14	9	-	No tumour	8	S	FACS, B/C
	-	-	-	8	S	FACS, B/C
17-12	-	-	-	8	M	B/C
28-14	9	-	No tumour	5	S	IHC, FACS, B/C
	9	-	No tumour	5	S	IHC, FACS, B/C
	-	-	-	5	S	H&E, FACS, B/C
	-	-	-	5	S	H&E, FACS, B/C
	-	-	-	5	S	FACS, B/C
	-	-	-	5	S	FACS, B/C
	-	-	-	5	S	FACS, B/C
	-	-	-	7	-	No tumour
21-13	9	-	No tumour	8	S	IHC, FACS, B/C
	-	-	-	9	-	No tumour
	-	-	-	9	-	No tumour
17-13	-	-	-	8	-	No tumour
13-12	9	-	No tumour	9	-	No tumour
42-14	-	-	-	9	-	No tumour
38-14	9	-	No tumour	-	-	-
33-14	9	-	No tumour	9	-	No tumour
	9	-	No tumour	-	-	-
	9	-	No tumour	-	-	-
15-14	9	-	No tumour	9	-	No tumour

The following information is indicated for each primary tumour tested for secondary tumour formation: the experimental conditions used to obtain the primary tumour, the size(s) of the secondary tumour(s) produced and the subsequent analyses performed, where applicable (B/C, barcode analysis). Small (S), medium (M), and large (L) refer to tumours for which the longest axis was measured as <5 mm, from 5 to 10 mm, or 10 to a maximum of 15 mm, respectively, to conform to our institutionally approved animal use protocols that stipulate tumours must be removed when they reach a size of 1 cm<sup>3</sup>. When an experiment was not performed or a data field was not applicable, this was indicated with '-'. When a transplant was performed and no tumour was produced, this is indicated as 'no tumour'. The subsequent analyses performed are indicated, where applicable (B/C, barcode analysis).

# FGF signalling regulates bone growth through autophagy

Laura Cinque<sup>1,2,\*</sup>, Alison Forrester<sup>1,2,3,\*</sup>, Rosa Bartolomeo<sup>1,2</sup>, Maria Svelto<sup>1,2,3</sup>, Rossella Venditti<sup>1</sup>, Sandro Montefusco<sup>1</sup>, Elena Polishchuk<sup>1</sup>, Edoardo Nusco<sup>1</sup>, Antonio Rossi<sup>4</sup>, Diego L. Medina<sup>1</sup>, Roman Polishchuk<sup>1</sup>, Maria Antonietta De Matteis<sup>1</sup> & Carmine Settembre<sup>1,2,3</sup>

**Skeletal growth relies on both biosynthetic and catabolic processes<sup>1,2</sup>. While the role of the former is clearly established, how the latter contributes to growth-promoting pathways is less understood. Macroautophagy, hereafter referred to as autophagy, is a catabolic process that plays a fundamental part in tissue homeostasis<sup>3</sup>. We investigated the role of autophagy during bone growth, which is mediated by chondrocyte rate of proliferation, hypertrophic differentiation and extracellular matrix (ECM) deposition in growth plates<sup>4</sup>. Here we show that autophagy is induced in growth-plate chondrocytes during post-natal development and regulates the secretion of type II collagen (Col2), the major component of cartilage ECM. Mice lacking the autophagy related gene 7 (*Atg7*) in chondrocytes experience endoplasmic reticulum storage of type II procollagen (PC2) and defective formation of the Col2 fibrillary network in the ECM. Surprisingly, post-natal induction of chondrocyte autophagy is mediated by the growth factor FGF18 through FGFR4 and JNK-dependent activation of the autophagy initiation complex VPS34–beclin-1. Autophagy is completely suppressed in growth plates from *Fgf18*<sup>-/-</sup> embryos, while *Fgf18*<sup>+/-</sup> heterozygous and *Fgfr4*<sup>-/-</sup> mice fail to induce autophagy during post-natal development and show decreased Col2 levels in the growth plate. Strikingly, the *Fgf18*<sup>+/-</sup> and *Fgfr4*<sup>-/-</sup> phenotypes can be rescued *in vivo* by pharmacological activation of autophagy, pointing to autophagy as a novel effector of FGF signalling in bone. These data demonstrate that autophagy is a developmentally regulated process necessary for bone growth, and identify FGF signalling as a crucial regulator of autophagy in chondrocytes.**

To understand the role of autophagy during bone growth, we analysed the femoral growth plates of transgenic mice that express the green fluorescent protein (GFP)-tagged autophagosome marker MAP1LC3 (GFP–LC3<sup>tg/+</sup>)<sup>5</sup>. We observed an increase in the number of autophagic vesicles (AVs) during post-natal bone development (Fig. 1a, quantification in Fig. 1b). We confirmed this observation by transmission electron microscopy (TEM) (Extended Data Fig. 1a) and by quantifying the conversion of LC3 (LC3I) to the autophagosome-associated lipidated form (LC3II)<sup>6</sup> in femoral growth plates of wild-type mice (Fig. 1c). Inhibition of lysosomal function by leupeptin administration further increased the levels of LC3II in the growth plate of post-natal day (P)6 but not P2 mice (Extended Data Fig. 1b). The age-dependent increase of LC3II suggests that autophagy may have a role during post-natal bone development. To determine if this was the case, we deleted *Atg7* in chondrocytes by crossing a mouse line carrying the *Atg7* floxed allele (*Atg7*<sup>fl/fl</sup>)<sup>7</sup> with the *Prx1*-Cre<sup>8</sup> and the *Col2a1*-Cre<sup>9</sup> lines, which express the Cre protein in the limb mesenchymal cells and in mature chondrocytes, respectively (Extended Data Fig. 1c–f). *Atg7*<sup>fl/fl</sup>; *Prx1*-Cre and *Atg7*<sup>fl/fl</sup>; *Col2a1*-Cre mice were born at the expected Mendelian

ratio, with bones of normal shapes and sizes (Extended Data Fig. 2a). Starting at P9 the *Atg7*<sup>fl/fl</sup>; *Prx1*-Cre mice showed reduced femoral and tibia lengths compared with control mice (Extended Data Fig. 2b). A similar, albeit milder, phenotype was also observed in *Atg7*<sup>fl/fl</sup>; *Col2a1*-Cre mice (Extended Data Fig. 2c, d). This difference could be due either to the residual autophagy observed in *Atg7*<sup>fl/fl</sup>; *Col2a1*-Cre growth plates (see Extended Data Fig. 1c) or to the contribution of additional cell populations targeted by the *Prx1*-Cre and not by the *Col2a1*-Cre line. Histological analyses of femoral and tibia growth plates from P6 and P9 *Atg7*<sup>fl/fl</sup>; *Prx1*-Cre and *Atg7*<sup>fl/fl</sup>; *Col2a1*-Cre mice showed preserved architecture and normal rates of chondrocyte differentiation, proliferation and terminal apoptosis (Extended Data Fig. 3a–e and data not shown). We then analysed the composition of growth-plate ECM. The levels of glycosaminoglycans were only slightly reduced in the growth plate of *Atg7*<sup>fl/fl</sup>; *Prx1*-Cre and *Atg7*<sup>fl/fl</sup>; *Col2a1*-Cre mice compared with controls (Extended Data Fig. 4a). PC2 is the main protein synthesized in chondrocytes, and Col2 constitutes the majority of cartilaginous ECM<sup>10</sup>. Col2 levels were normal in the growth plates of *Atg7*<sup>fl/fl</sup>; *Prx1*-Cre and *Atg7*<sup>fl/fl</sup>; *Col2a1*-Cre mice at birth, but were severely reduced at P9 compared to control mice (Fig. 1d, e and Extended Data Fig. 4b). Consistently, TEM microscopy showed sparse and disorganized Col2 fibril networks in femoral growth plates of *Atg7*<sup>fl/fl</sup>; *Prx1*-Cre mice compared with controls (Fig. 1f).

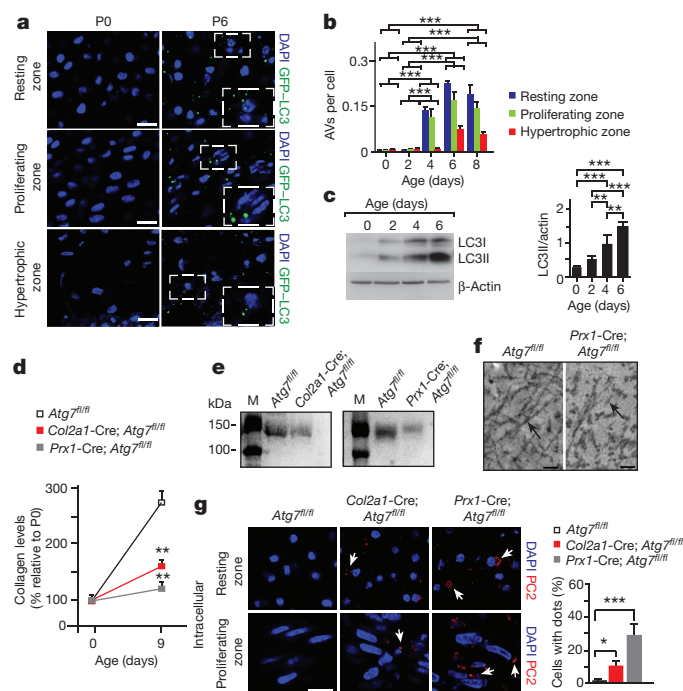
We detected ectopic accumulation of PC2 in the endoplasmic reticulum (ER) of chondrocytes lacking autophagy (Fig. 1g and Extended Data Fig. 4c). Consistently, TEM analysis showed that the ER cisternae of *Atg7*<sup>fl/fl</sup>; *Prx1*-Cre chondrocytes were enlarged and filled with electron-dense material (Extended Data Fig. 4d). Inhibition of autophagy with Spautin-1 (ref. 11) or with RNA interference targeting *Atg7* (*Atg7* knockdown) in cultured Swarm rat chondrosarcoma chondrocytes (RCS) led to defective PC2 secretion and to the retention of PC2 in the ER (Fig. 2a–c and Extended Data Fig. 5a, b).

These data suggest that chondrocyte autophagy regulates post-natal bone growth in part by controlling the levels of Col2 secreted in the ECM.

We observed that a fraction of PC2 co-localizes with GFP–LC3 puncta, suggesting that PC2 is an autophagy substrate in chondrocytes (Fig. 2d and Extended Data Fig. 5c–e). Consistently, dual-colour (mCherry–PC2 and GFP–LC3) live-cell imaging experiments showed sequestration of PC2 by GFP–LC3-positive vesicles (Fig. 2e, f and Supplementary Video 1). The collagen-specific chaperone HSP47, which associates with native PC2 triple helices in the ER<sup>12</sup>, was excluded from the AVs containing PC2, suggesting that autophagy sequesters non-native PC2 molecules in the ER (Extended Data Fig. 5f). Whereas in control chondrocytes HSP47 showed a diffuse distribution and normal ER–Golgi trafficking, in chondrocytes with

<sup>1</sup>Telethon Institute of Genetics and Medicine (TIGEM), Via Campi Flegrei, 34, 80078 Pozzuoli (NA), Italy. <sup>2</sup>Dulbecco Telethon Institute, Via Campi Flegrei, 34, 80078 Pozzuoli (NA), Italy. <sup>3</sup>Medical Genetics Unit, Department of Medical and Translational Science, Federico II University, Via Pansini 5, 80131 Naples, Italy. <sup>4</sup>Department of Molecular Medicine, Biochemistry Unit, University of Pavia, 27100 Pavia, Italy.

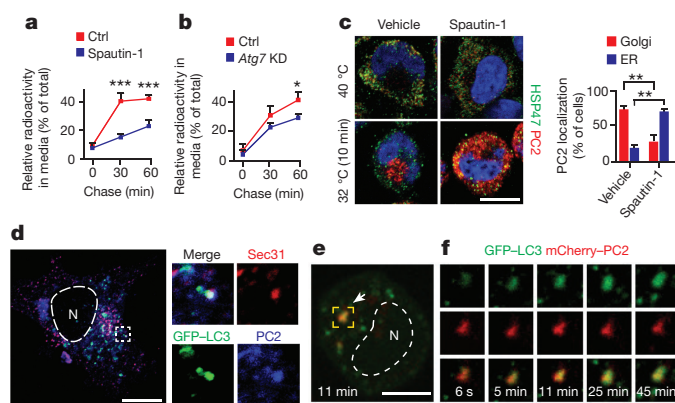
\*These authors contributed equally to this work.



impaired autophagy it was clustered with PC2 aggregates (Extended Data Fig. 6a–c). These data suggest that autophagy regulates PC2 homeostasis in the ER, consistent with the recently described role of autophagy as an ER quality control pathway<sup>13–15</sup> (Extended Data Fig. 6d).

During autophagy, AVs target their cargo to lysosomes for degradation. PC2 transport assays using RCS chondrocytes co-expressing mCherry-PC2 and GFP-LAMP1 showed progressive and autophagy-dependent accumulation of PC2 in GFP-LAMP1 vesicles (Extended Data Fig. 6e, f). By using two independent approaches (total internal reflection fluorescence (TIRF) microscopy and tannic acid treatment) we failed to detect LC3- or LAMP1-positive vesicles fusing with the plasma membrane, suggesting that neither AVs nor lysosomes mediate PC2 exocytosis (Supplementary Videos 2, 3 and Extended Data Fig. 6g, h).

Next we tested the hypothesis that autophagy might be co-regulated with Col2 production during bone development. We investigated the role of FGF18, a regulator of Col2 production in chondrocytes<sup>16</sup>, in

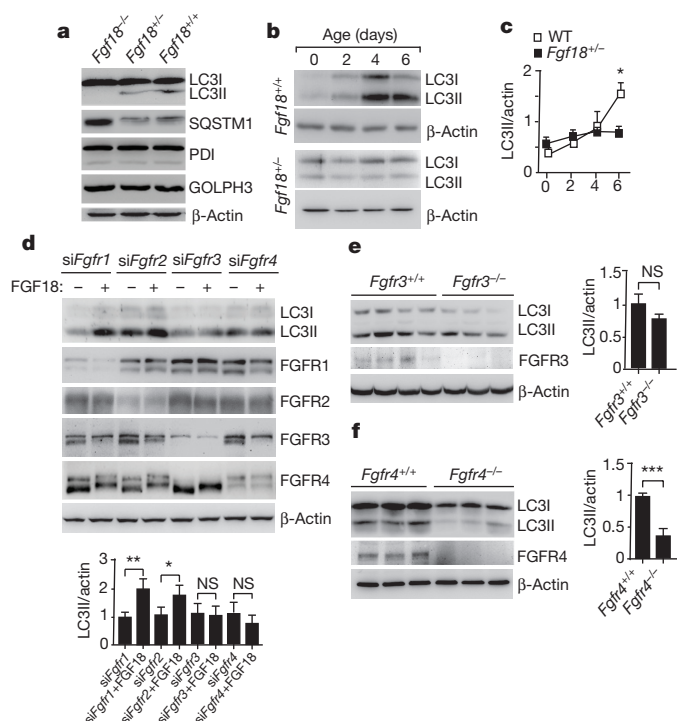


autophagy. Primary chondrocytes isolated from GFP-LC3 mice were stimulated with FGF18 and other chondrogenic factors<sup>17</sup> and AV biogenesis was assessed (Extended Data Fig. 7a). Among the factors tested, only FGF18 was able to significantly increase AV number (Extended Data Fig. 7a, b). FGF18 enhanced the autophagic flux also in wild-type primary chondrocytes (Extended Data Fig. 7c), and in RCS chondrocytes expressing the tandem fluorescently-tagged LC3 (monomeric red fluorescent protein (mRFP)-enhanced (e)GFP-LC3) protein<sup>18</sup> (Extended Data Fig. 7d). The growth plates of *Fgf18*<sup>-/-</sup> embryonic day (E)18.5 embryos were characterized by undetectable levels of LC3II and accumulation of the autophagy receptor SQSTM1 (also known as P62) compared with controls (Fig. 3a). The levels of other organelle markers, such as PDI (ER) and GOLPH3 (Golgi) were not affected (Fig. 3a). *Fgf18*<sup>-/-</sup> mice exhibit neonatal lethality<sup>19</sup>; therefore, we analysed the growth plates of *Fgf18*<sup>+/-</sup> mice. We observed defective post-natal induction of autophagy and reduced AV biogenesis in *Fgf18*<sup>+/-</sup> chondrocytes compared with control mice (Fig. 3b, c and Extended Data Fig. 7e, f). Consistently, SQSTM1 levels were higher in *Fgf18*<sup>+/-</sup> growth plates compared with controls at P30 (Extended Data Fig. 7g). Thus, FGF18 is a critical regulator of chondrocyte autophagy during skeletal development.

RNA interference of either *Fgfr3* or *Fgfr4*, but not of *Fgfr1* and *Fgfr2*, inhibits FGF18-induced autophagy in RCS chondrocytes (Fig. 3d and Extended Data Fig. 8a). Stimulation of chondrocytes with FGF18 induced autophosphorylation of both FGFR3 and FGFR4 (Extended Data Fig. 8b). *In vivo*, growth-plate chondrocytes express both FGFR3 and FGFR4 (Extended Data Fig. 8c). However, we only observed a significant decrease of LC3II levels in the growth plates of *Fgfr4*<sup>-/-</sup> mice (Fig. 3e, f). These data suggest that the regulation of autophagy by FGF18 is mainly mediated through FGFR4.

Canonical GGF signalling activates the mitogen-activated protein kinase (MAPK) pathway. The growth plates of *Fgf18*<sup>+/-</sup> mice showed lower levels of JNK1/2 kinase activation than control mice

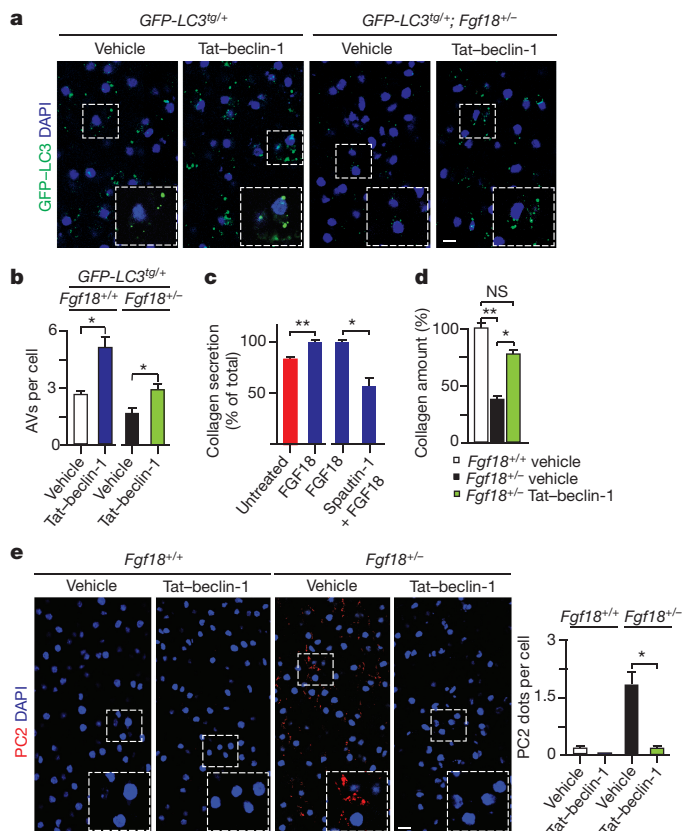




**Figure 3 | FGF18 regulates autophagy in growth-plate chondrocytes.** **a**, Western blot analysis of LC3I/II, SQSTM1, PDI and GOLPH3 in femoral growth plates of E18.5 mice with the indicated genotypes.  $\beta$ -Actin was used as a loading control. Blot is representative of 3 independent experiments. **b**, Western blot analysis of LC3I/II of femoral growth plates from control and *Fgf18*<sup>+/-</sup> mice at the indicated ages.  $\beta$ -Actin was used as a loading control. **c**, Quantification of LC3II levels. WT, wild type. Values are expressed as mean ( $\pm$  s.e.m.) of  $n = 3$  per mice for each time point. ANOVA,  $P = 9.04 \times 10^{-5}$ ; Tukey's post-hoc test,  $*P < 0.05$ . **d**, Western blot analysis of LC3II levels in RCS chondrocytes treated with short interfering RNA (siRNA) for *Fgfr1*, *Fgfr2*, *Fgfr3* or *Fgfr4* and then stimulated with FGF18 for 2 h. Bafilomycin A1 (BafA1) was added (200 nM, 3 h). Bar graphs represent mean values  $\pm$  s.e.m. of  $n = 6$  independent experiments. Student's *t*-test (FGF18-treated versus untreated),  $*P < 0.05$ ,  $**P < 0.005$ . NS, not significant. **e**, **f**, Western blot analysis of LC3I/II, FGFR3 (**e**) and FGFR4 (**f**) in femoral growth plates from control, *Fgfr3*<sup>-/-</sup> and *Fgfr4*<sup>-/-</sup> littermate mice ( $n = 3$  per genotype) at P6. Graph shows mean ( $\pm$  s.e.m.) of LC3II levels relative to  $\beta$ -actin. Student's *t*-test,  $***P < 0.0005$ . NS, not significant.  $n = 3$  mice per genotype.

(Extended Data Fig. 8d). No changes were observed in the activation states of other members of the MAPK pathway (ERK and P38) or of other kinases involved in autophagy (Extended Data Fig. 8e). Active JNK1 phosphorylates BCL2 and disrupts the BCL2–beclin-1 complex<sup>20</sup>, leading to the activation of the class III phosphatidylinositol-3-OH kinase (PtdIns(3)K) VPS34–beclin-1 complex<sup>21</sup>. Consistently, we observed that: (1) FGF18 increased the phosphorylation of BCL2 in a JNK-dependent manner (Extended Data Fig. 9a); (2) FGF18 decreased the interaction of BCL2–beclin-1 (Extended Data Fig. 9b); (3) FGF18 increased VPS34–beclin-1 complex activity in a JNK-dependent manner, as indicated by the amount of phosphatidylinositol-3-phosphate (PtdIns(3)P) levels produced (Extended Data Fig. 9c, d); and (4) enhancing beclin-1 activity by intraperitoneal injection of a synthetic beclin-1 activator (Tat-beclin-1) peptide<sup>22</sup> normalized autophagy levels in the growth plates of *Fgf18*<sup>+/-</sup>; GFP-LC3<sup>tg/+</sup> mice (Fig. 4a, quantification in Fig. 4b). Thus, FGF18 induces autophagy, at least in part, through the regulation of VPS34–beclin-1 complex activity.

Next we asked whether FGF18 controls Col2 secretion through modulation of autophagy. RCS chondrocytes stimulated with FGF18 exhibited higher efficiency of PC2 secretion compared with non-stimulated cells, but addition of the autophagy inhibitor Spautin-1 hampered this



**Figure 4 | FGF18 regulates Col2 secretion via autophagy.**

**a**, Representative images of GFP-LC3 puncta (autophagosomes) in femoral growth plates from *Fgf18*<sup>+/-</sup>; GFP-LC3<sup>tg/+</sup> and *Fgf18*<sup>+/-</sup>; GFP-LC3<sup>tg/+</sup> mice at P6. Tat-beclin-1 peptide was administered where indicated (2 mg kg<sup>-1</sup>, daily for 6 days). The insets show a higher magnification of selected areas. Scale bar, 10  $\mu$ m. **b**, Quantification of data. Values are mean  $\pm$  s.e.m. of  $n = 3$  mice per group. Two sections per mouse ( $N = 200$ ) were analysed. Student's *t*-test,  $*P < 0.05$ . **c**, PC2 secretion in untreated versus FGF18-treated (25 ng ml<sup>-1</sup>) and in FGF18-treated versus FGF18 plus Spautin-1-treated (50  $\mu$ M) RCS chondrocytes. Values (means  $\pm$  s.e.m.;  $n = 3$  independent experiments) were expressed as percentage relative to FGF18 treatment. Student's *t*-test,  $*P < 0.05$ ,  $**P < 0.005$ . **d**, Total collagen in femoral growth plates of *Fgf18*<sup>+/-</sup> and *Fgf18*<sup>+/-</sup> mice at P9 treated with vehicle or with Tat-beclin-1 (2 mg kg<sup>-1</sup>, daily for 9 days). Mean  $\pm$  s.e.m. of 4 mice per group. ANOVA,  $P = 0.0049$ ; Tukey's post-hoc test,  $*P < 0.05$ ,  $**P < 0.005$ ; NS, not significant. **e**, Confocal analysis of intracellular PC2 in *Fgf18*<sup>+/-</sup> and *Fgf18*<sup>+/-</sup> growth plates at P6. Tat-beclin-1 was administered where indicated. The insets show a high magnification of selected areas. Red, collagen; blue, DAPI. Scale bar, 10  $\mu$ m. Graph shows quantification of data. Mean  $\pm$  s.e.m. of  $n = 4$  mice per group. Three sections per mice containing 100 cells were analysed. ANOVA,  $P = 0.0091$ ; Tukey's post-hoc test,  $*P < 0.05$ .

increase (Fig. 4c). The *Fgf18*<sup>+/-</sup> growth plates were characterized by a severe reduction of collagen levels in the ECM and by the presence of intracellular PC2 deposits in chondrocytes compared with control mice (Fig. 4d, e). Notably, the few GFP-labelled AVs detectable in the growth plates of *Fgf18*<sup>+/-</sup>; GFP-LC3<sup>tg/+</sup> mice contained PC2, further demonstrating that PC2 is an autophagy substrate *in vivo* (Extended Data Fig. 9e). Tat-beclin-1 treatment restored Col2 levels in the growth plates of *Fgf18*<sup>+/-</sup> mice and completely eliminated the intracellular accumulation of PC2 in *Fgf18*<sup>+/-</sup> chondrocytes (Fig. 4d, e and Extended Data Fig. 9f). In addition, Tat-beclin-1 treatment increased Col2 levels and femur sizes in P9 and P15 *Fgfr4*<sup>-/-</sup> mice (Extended Data Fig. 9g–i). These data indicate that FGF signalling regulates the composition of growth-plate ECM through modulation of autophagy. Future studies will be required to investigate in detail the bone phenotype in *Fgfr4*<sup>-/-</sup> mice.

We have shown a previously unanticipated<sup>2,23–25</sup> role of chondrocyte autophagy in bone growth (Extended Data Fig. 10). In addition, by showing that chondrocyte autophagy is regulated by FGF signalling, we demonstrated that autophagy is a developmentally regulated process. The FGF pathway has been extensively studied in the bone due to its involvement in many forms of skeletal disorders<sup>26,27</sup>. A genome-wide association study showed that *FGFR4* sequence variations may influence human height<sup>28</sup>. Our study describes a new mechanism by which FGF signalling regulates post-natal bone growth and identifies autophagy as a critical regulator of vertebrate development.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 11 September 2014; accepted 7 October 2015.**

**Published online 23 November 2015.**

- Karsenty, G. & Wagner, E. F. Reaching a genetic and molecular understanding of skeletal development. *Dev. Cell* **2**, 389–406 (2002).
- Shapiro, I. M., Layfield, R., Lotz, M., Settembre, C. & Whitehouse, C. B. B. On autophagy: the role of autophagy in skeletal biology. *Autophagy* **10**, 7–19 (2014).
- Mizushima, N. & Komatsu, M. Autophagy: renovation of cells and tissues. *Cell* **147**, 728–741 (2011).
- Wilsman, N. J., Farnum, C. E., Leiferman, E. M., Fry, M. & Barreto, C. Differential growth by growth plates as a function of multiple parameters of chondrocytic kinetics. *J. Orthop. Res.* **14**, 927–936 (1996).
- Mizushima, N., Yamamoto, A., Matsui, M., Yoshimori, T. & Ohsumi, Y. *In vivo* analysis of autophagy in response to nutrient starvation using transgenic mice expressing a fluorescent autophagosome marker. *Mol. Biol. Cell* **15**, 1101–1111 (2004).
- Kabeya, Y. *et al.* LC3, a mammalian homologue of yeast Apg8p, is localized in autophagosome membranes after processing. *EMBO J.* **19**, 5720–5728 (2000).
- Komatsu, M. *et al.* Impairment of starvation-induced and constitutive autophagy in Atg7-deficient mice. *J. Cell Biol.* **169**, 425–434 (2005).
- Logan, M. *et al.* Expression of Cre recombinase in the developing mouse limb bud driven by a *Prrx1* enhancer. *Genesis* **33**, 77–80 (2002).
- Ovchinnikov, D. A., Deng, J. M., Ogunrinu, G. & Behringer, R. R. Col2a1-directed expression of Cre recombinase in differentiating chondrocytes in transgenic mice. *Genesis* **26**, 145–146 (2000).
- Olsen, B. R., Reginato, A. M. & Wang, W. Bone development. *Annu. Rev. Cell Dev. Biol.* **16**, 191–220 (2000).
- Liu, J. *et al.* Beclin1 controls the levels of p53 by regulating the deubiquitination activity of USP10 and USP13. *Cell* **147**, 223–234 (2011).
- Ishida, Y. & Nagata, K. Hsp47 as a collagen-specific molecular chaperone. *Methods Enzymol.* **499**, 167–182 (2011).
- Ishida, Y. *et al.* Autophagic elimination of misfolded procollagen aggregates in the endoplasmic reticulum as a means of cell protection. *Mol. Biol. Cell* **20**, 2744–2754 (2009).
- Teckman, J. H. & Perlmutter, D. H. Retention of mutant  $\alpha_1$ -antitrypsin Z in endoplasmic reticulum is associated with an autophagic response. *Am. J. Physiol. Gastrointest. Liver Physiol.* **279**, G961–G974 (2000).
- Houck, S. A. *et al.* Quality control autophagy degrades soluble ERAD-resistant conformers of the misfolded membrane protein GnRHR. *Mol. Cell* **54**, 166–179 (2014).
- Moore, E. E. *et al.* Fibroblast growth factor-18 stimulates chondrogenesis and cartilage repair in a rat model of injury-induced osteoarthritis. *Osteoarthritis Cartilage* **13**, 623–631 (2005).
- Karsenty, G., Kronenberg, H. M. & Settembre, C. Genetic control of bone formation. *Annu. Rev. Cell Dev. Biol.* **25**, 629–648 (2009).
- Kimura, S., Fujita, N., Noda, T. & Yoshimori, T. Monitoring autophagy in mammalian cultured cells through the dynamics of LC3. *Methods Enzymol.* **452**, 1–12 (2009).
- Liu, Z., Xu, J., Colvin, J. S. & Ornitz, D. M. Coordination of chondrogenesis and osteogenesis by fibroblast growth factor 18. *Genes Dev.* **16**, 859–869 (2002).
- Wei, Y., Pattingre, S., Sinha, S., Bassik, M. & Levine, B. JNK1-mediated phosphorylation of Bcl-2 regulates starvation-induced autophagy. *Mol. Cell* **30**, 678–688 (2008).
- Liang, X. H. *et al.* Induction of autophagy and inhibition of tumorigenesis by beclin 1. *Nature* **402**, 672–676 (1999).
- Shoji-Kawata, S. *et al.* Identification of a candidate therapeutic autophagy-inducing peptide. *Nature* **494**, 201–206 (2013).
- Lotz, M. K. & Caramés, B. Autophagy and cartilage homeostasis mechanisms in joint health, aging and OA. *Nature Rev. Rheumatol.* **7**, 579–587 (2011).
- Srinivas, V., Bohensky, J., Zahm, A. M. & Shapiro, I. M. Autophagy in mineralizing tissues: microenvironmental perspectives. *Cell Cycle* **8**, 391–393 (2009).
- Vuppapalapati, K. K. *et al.* Targeted deletion of autophagy genes *Atg5* or *Atg7* in the chondrocytes promotes caspase-dependent cell death and leads to mild growth retardation. *J. Bone Miner. Res.* <http://dx.doi.org/10.1002/jbmr.2575> (2015).
- Ornitz, D. M. & Marie, P. J. FGF signaling pathways in endochondral and intramembranous bone development and human genetic disease. *Genes Dev.* **16**, 1446–1465 (2002).
- Settembre, C. *et al.* Proteoglycan desulfation determines the efficiency of chondrocyte autophagy and the extent of FGF signaling during endochondral ossification. *Genes Dev.* **22**, 2645–2650 (2008).
- Lango Allen, H. *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832–838 (2010).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank A. Ballabio, G. Karsenty, A. Auricchio, M. Sandri and G. Diez Roux for critical reading of the manuscript. We thank L. Staiano and C. Lanzara for technical support. We thank A. Carissimo for statistical analysis. We thank N. Mizushima, D. Ornitz and J. Seavitt for sharing mouse lines used in this work. This work was supported by grants from the Italian Telethon Foundation (TCP12008), Marie Curie European Reintegration Grant (PCIG13-GA-2013-618805), STAR/Banco di Napoli and the Italian Ministry of Research (FIRB RBFR13LH4X) to C.S.

**Author Contributions** L.C. performed the majority of the experiments involving mice. A.F. performed the majority of the *in vitro* experiments. R.B. and M.S. performed immunoprecipitation experiments. R.V. provided essential reagents and helped with the collagen experiments. S.M. performed the high-content screening experiments. E.P. performed TEM analysis. E.N. performed Tat-beclin-1 injections and tissue harvesting. A.R. provided critical suggestions and protocols. R.P. supervised the TEM. D.L.M. supervised the high-content screening experiments. M.A.D.M. provided critical suggestions, protocols and reagents. C.S. conceived, designed and supervised the study and wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.S. ([settembre@tigem.it](mailto:settembre@tigem.it)).

## METHODS

**Animals.** The *Atg7<sup>fl/fl</sup>* line was from T. Eissa; the GFP-LC3 mouse line was from N. Mizushima. The *Prx1*-Cre line was purchased from Jackson Laboratories (strain no. 005584). The *Col2a1*-Cre line was from B. Lee. The *Fgf18*- and *Fgfr3*-knockout mice were from D. Ornitz. The *Fgfr4*-knockout mice were from J. Seavitt. All mice used were maintained in a C57BL/6 strain background. The number of mice used in each experiment is specified in figure legends. Sex of the mice was not taken into account until P9. At later time points (P30 and P120) only male mice were analysed. Mice were randomly assigned to treatment groups. The investigators were not blinded to allocation during experiments and outcome assessment.

Experiments were conducted in accordance with the guidelines of the Animal Care and Use Committee of Cardarelli Hospital in Naples and authorized by the Italian Ministry of Health.

**Skeletal staining.** Skeletons were fixed in 95% ethanol overnight and stained with alcian blue and alizarin red according to standardized protocols (<http://empres.har.mrc.ac.uk/browser/>). Measurement of bone length was performed using ImageJ software.

**Tissue histology, immunohistochemistry and immunofluorescence.** Histology was performed according to standardized procedures (<http://empres.har.mrc.ac.uk/browser/>). Briefly, femurs were fixed overnight in 4% (wt/vol) paraformaldehyde (PFA) and then demineralized in 10% EDTA (pH 7.4) for 48 h (demineralization was performed only if specimens were isolated from mice older than P5). Specimens were then dehydrated, embedded in paraffin and sectioned at 7  $\mu$ m, and stained with haematoxylin and eosin. For BrdU staining, mice were injected with 100  $\mu$ l of 10 mM BrdU (Sigma) 4 h before being killed. BrdU incorporation was detected using a Zymed BrdU staining kit (Invitrogen). The TdT-mediated dUTP nick end labelling (TUNEL) assay was performed using the *In situ* Cell Death Detection kit (Roche). Counterstaining was performed using haematoxylin.

For immunofluorescence, femurs were dissected from euthanized mice and fixed with buffered 4% PFA overnight at 4°C, then washed with PBS and cryoprotected in successive sucrose solutions diluted with PBS (10% for 2 h, 20% for several hours and 30% overnight at 4°C; all wt/vol), and finally embedded in OCT (Sakura). Cryostat sections were cut at 10  $\mu$ m. Sections were blocked and permeabilized in 3% (wt/vol) BSA, 5% fetal bovine serum (FBS) in PBS plus 0.3% Triton X-100 for 3 h and then incubated with the primary antibody overnight. Sections were washed three times with 3% BSA in PBS plus 0.3% Triton X-100 and then incubated for 3 h with secondary antibodies conjugated with Alexa Fluor 488, or Alexa Fluor 568.

The extracellular Col2 staining was performed by pre-treating sections with chondroitinase ABC (Sigma) at 0.2 U ml<sup>-1</sup> for 1 h at 37°C before the blocking step. Intracellular PC2 staining was performed without chondroitinase ABC pre-treatment to stain only the PC2 molecules that were not masked by proteoglycans.

Primary antibodies used were: GFP, LAMP1 and HSP47 (Abcam), Col2a1 (1:30, Hybridoma Bank, II6B3), FGFR3 and FGFR4 (Cell Signaling), VapA, Sec31, giantin, GM130, P115 and calreticulin were previously described<sup>29</sup>. Nuclei were stained with DAPI and sections were mounted with vectashield (Vector laboratories). Images were captured using a Zeiss LSM700 confocal microscope. Co-localization analysis was performed calculating Mander's coefficient using ImageJ (co-localization analysis plug-in). Semi-quantitative analyses were performed by an investigator blinded to the genotype of the mice.

**Collagen quantification and analysis.** *Colorimetric assay.* This was performed using the Sircol soluble collagen assay (Biocolor) following the manufacturer's protocol. Briefly, femoral and tibia cartilages were microdissected and collagen was acid pepsin extracted and complexed with Sircol dye. Absorbance was measured at 555 nm and concentration was calculated using a standard curve. Values were normalized to DNA levels calculated measuring the absorbance at 260 nm.

*Electrophoretic analysis.* Femoral cartilages were isolated from three mice with the same genotype, pooled and homogenized in 0.5 ml of 1 mg ml<sup>-1</sup> cold (4°C) pepsin in 0.2 M NaCl, 0.5 M acetic acid to pH 2.1 with HCl and then digested at 4°C for 24 h, twice. The pellet was discarded and an equal volume (1 ml) of 4 M NaCl in 1 M acetic acid was added to precipitate collagen. The pellet was then resuspended in 0.8 ml of 0.2 M NaCl in 0.5 M acetic acid and was precipitated again three times. After the last precipitation the pellet was washed twice with 70% EtOH to remove residual NaCl. The pellet was then dissolved in 0.8 ml 0.5 M acetic acid, and lyophilized. Subsequently it was resuspended in Laemmli buffer without EtSH at a concentration of 2 mg ml<sup>-1</sup>, denatured at 80°C for 5 min and loaded on 6% SDS-PAGE. Gels were then stained with Coomassie blue R-250.

**GAG quantification.** GAG (glycosaminoglycan) quantification was performed using the Blyscan sulfated glycosaminoglycan assay (Biocolor) following the manufacturer's protocol. Briefly, femoral and tibia cartilages were microdissected and GAGs were papain extracted at 65°C overnight and complexed with Blyscan dye. Absorbance was measured at 656 nm and concentration was calculated using a

standard curve. Values were normalized to DNA levels calculated measuring the absorbance at 260 nm.

**Transmission electron microscopy.** For EM analysis, growth plates were fixed in 1% glutaraldehyde in 0.2 M HEPES buffer. Small blocks of growth plates were then post-fixed in uranyl acetate and in OsO<sub>4</sub>. After dehydration through a graded series of ethanol, tissue samples were cleared in propylene oxide, embedded in Epoxy resin (Epon 812) and polymerized at 60°C for 72 h. From each sample, thin sections were cut with a Leica EM UC6 ultramicrotome and images were acquired using a FEI Tecnai – 12 (FEI) electron microscope equipped with Veletta CCD camera for digital image acquisition.

**Tat-beclin-1 peptide and Leupeptin treatment.** Newborn mice were intraperitoneally injected daily with Tat-beclin-1 peptide (Beclin-1 Activator II, retro-inverso Tat-beclin-1, Millipore) at 2 mg kg<sup>-1</sup> resuspended in PBS. Control mice were injected with vehicle only. Mice were killed after 6 or 9 days for PC2 immunofluorescence experiments and total collagen quantification, respectively.

Leupeptin (Sigma catalogue L2884) was resuspended in water at 10 mM. Mice were given an intraperitoneal injection at 40 mg kg<sup>-1</sup>. Six hours after injection tissues were harvested and processed for western blotting.

**Tissue protein extracts for western blotting.** Femoral and tibia cartilages were microdissected and lysed using a TissueLysr (Qiagen) in RIPA buffer supplemented with 0.5% SDS, PhosSTOP and EDTA-free protease inhibitor tablets (Roche). Samples were incubated for 30 min on ice, briefly sonicated on ice and the soluble fraction was isolated by centrifugation at 14,000 r.p.m. for 10 min at 4°C.

**Chemicals.** FGF18 (50 ng ml<sup>-1</sup> unless otherwise indicated), PTHrP (10  $\mu$ g ml<sup>-1</sup>) and BMP2 (500 ng ml<sup>-1</sup>) were from Peprotech, rhSHH (10  $\mu$ g ml<sup>-1</sup>) was from R&D Systems. JNK inhibitor (SP600125, Sigma-Aldrich) was used at 50  $\mu$ M. Tannic acid (Fluka Chemika) was used at 0.5% final concentration in the medium for 1 h at 37°C. Bafilomycin A1 (Sigma) was used at 200 nM.

**Cell culture, transfections, siRNA and plasmids.** Primary cultured chondrocytes were prepared from rib cartilage of P5 mice. Rib cages were first incubated in DMEM (Euroclone) using 0.2% collagenase D (Roche) and after adherent connective tissue had been removed (1.5 h) the specimens were washed and incubated in fresh collagenase D solution for a further 4.5 h. Isolated chondrocytes were maintained in DMEM supplemented with 10% FCS (Invitrogen) and ascorbic acid (Sigma Aldrich) (50  $\mu$ g ml<sup>-1</sup>). Given the incomplete deletion of the *Atg7* gene in *Atg7<sup>fl/fl</sup>*; *Col2a1*-Cre chondrocytes (Extended Data Fig. 1c) and the lack of Cre expression in chondrocostal chondrocytes of *Prx1*-Cre mice, collagen secretion experiments were performed in the RCS chondrocyte cell line<sup>30</sup>, in which autophagy was inhibited by *Atg7* RNA interference and by Spautin-1 treatment. Cells were transfected with Lipofectamine LTX and Plus reagent (Invitrogen) following a reverse transfection protocol. For siRNA experiments, siGENOME SMARTpool siRNAs (Dharmacon Thermo Scientific) were transfected to the final concentration of 50 nM. Cells were harvested 72 h after transfection. Plasmids: GFP-LC3 and mRFP-GFP-LC3 were from T. Yoshimori, GFP-LAMP1 was from A. Fraldi mCherry-PC2 was previously described<sup>29</sup>; BCL2-haemagglutinin (HA) was from M. Renna, 2 × FYFE-GFP was from S. Tooze. FGFR3 plasmid was from Addgene and FGFR4 from DNASU plasmid repositories.

**Cell immunofluorescence.** Chondrocytes were fixed for 10 min in 4% PFA in PBS and permeabilized for 30 min in 0.05% (w/v) saponin, 0.5% (w/v) BSA, 50 mM NH<sub>4</sub>Cl and 0.02% Na<sub>3</sub>N in PBS (blocking buffer). For the detection of endogenous LC3 cells were methanol-fixed. The cells were incubated for 1 h with the primary antibodies, washed three times in PBS, incubated for 1 h with the secondary (Alexa Fluor-labelled) antibody, washed three times in PBS, incubated for 20 min with 1  $\mu$ g ml<sup>-1</sup> Hoechst 33342 and finally mounted in Mowiol. All confocal experiments showing co-localization were acquired using slice thickness of 0.5  $\mu$ m using the LSM 710 confocal microscope equipped with a 63 × 1.4 numerical aperture oil objective.

**Tannic acid treatment.** To prevent fusion of exocytic organelles with PM, chondrocytes were incubated for 1 h at 37°C in 20 mM HEPES-buffered DMEM supplemented with 0.5% tannic acid (TA). At the end of the incubation, TA-containing medium was removed and cells were fixed and processed for immunofluorescence.

**Live-cell imaging.** *Time-lapse microscopy.* RCS chondrocytes were reverse transfected with mCherry-PC2 and GFP-LC3 plasmids and plated in Mattek glass bottomed dishes. Collagen synchronization was performed by incubating cells at 40°C on the heated stage for 2.5 h. Collagen release was initiated by lowering the temperature of the stage to 32°C and medium being supplemented with 50  $\mu$ g ml<sup>-1</sup> of ascorbate, in a humidified atmosphere with 5% CO<sub>2</sub>. Time-lapse videos were acquired for 45 min, starting from collagen release, during temperature drop (4 min), and at 32°C. One frame was acquired roughly every 20 s with lasers set at 30% power or below.

**TIRF.** RCS chondrocytes were reverse transfected and plated in Mattek glass-bottomed dishes. RCS cells were synchronized on the heated stage for 2.5 h at 40°C and released at 32°C, in medium supplemented with 50  $\mu$ g ml<sup>-1</sup> ascorbic acid in



a humidified atmosphere with 5% CO<sub>2</sub>. The critical angle used was 65°, giving an evanescent field of 137 nm. Appropriate filter sets were used for GFP and mCherry detection. Frames were acquired on loop with no time delay (one frame roughly every 3 s), for 15 min.

All live-cell imaging experiments were performed with a ×60 Plan Apo oil immersion lens using a Nikon Eclipse Ti Spinning Disk microscope, and images and videos were annotated using the NIS Elements 4.20 software.

**Western blotting.** Cells were washed twice with PBS and then scraped in lysis buffer (RIPA lysis buffer in the presence of PhosSTOP and EDTA-free protease inhibitor tablets; Roche). Cell lysates were incubated on ice for 20 min, then the soluble fraction was isolated by centrifugation at 14,000 r.p.m. for 10 min at 4°C. Total protein concentration in cellular extracts was measured using the colorimetric BCA protein assay kit (Pierce Chemical). Protein extracts, separated by SDS-PAGE and transferred onto PVDF or nitrocellulose (for collagen) membranes, were probed with antibodies against FGFR3, FGFR4, phospho (p)-JNK, JNK, p-BCL2, p-c-JUN, PDI, p-P38, P38, beclin-1, p-ERK, ERK1/2, p-P70S6K, P70S6K, p-4EBP1, 4EBP1, ATG7, p-AKT, AKT, p-mTORC1, mTORC1 (Cell Signaling Technology), HA, H3 histone, VPS34 (Sigma-Aldrich), LC3B, β-actin (Novus Biologicals), SQSTM1 (BD Transduction Laboratories and Abnova), GOLPH3 (Abcam), GAPDH, p-AMPKα, AMPKα (Santa Cruz Biotechnology), type II collagen (CIIC1b; Hybridoma Bank). Proteins of interest were detected with horseradish peroxidase (HRP)-conjugated goat anti-mouse or anti-rabbit IgG antibody (1:2,000, Vector Laboratories) and visualized with the Super Signal West Dura substrate (Thermo Scientific), according to the manufacturer's protocol. The western blotting images were acquired using the Chemidoc-It imaging system (UVP) and band intensity was calculated using ImageJ software using the 'Gels and Plot lanes' plug-in.

**High-content screening analysis in GFP-LC3 primary chondrocytes.** Primary chondrocytes were plated in CellCarrier-96 Black plates (6005558, Perkin Elmer). After identifying the nuclei with Hoechst 33342 (405 nm) staining, a cytoplasmic mask was drawn using Col2 staining (568 nm). To carry out the analysis, the number of GFP-LC3 spots in the cytoplasm of Col2-positive cells were counted, and expressed per cell. Levels of co-localization between GFP-LC3 and Col2a1 were assessed and expressed as a percentage using the parameters: area of co-localization of red spots with area of green spots normalized to total area of green spots. Image acquisition was performed using Opera High Content Screening System (PerkinElmer); image analysis was performed using Acapella High Content Imaging and Analysis Software (PerkinElmer). Repeated-measures ANOVA was performed with Tukey's post-hoc test.

**Co-immunoprecipitation.** RCS chondrocytes (100-mm dish) were grown in DMEM medium with 10% FBS and antibiotics. For FGF18 treatment, 70–80% confluent cells were cultured overnight in DMEM with 10% adult bovine serum (Sigma-Aldrich) and then treated with FGF18 (50 ng ml<sup>-1</sup>, 2 h) (Peprotech) or dimethylsulfoxide (DMSO) vehicle. RCS chondrocytes were rinsed off the plate with ice-cold PBS, washed, and then scraped in IP lysis buffer (150 mM NaCl, 50 mM Tris-HCl pH 8.0, 1% NP-40, with one PhosSTOP and one EDTA-free

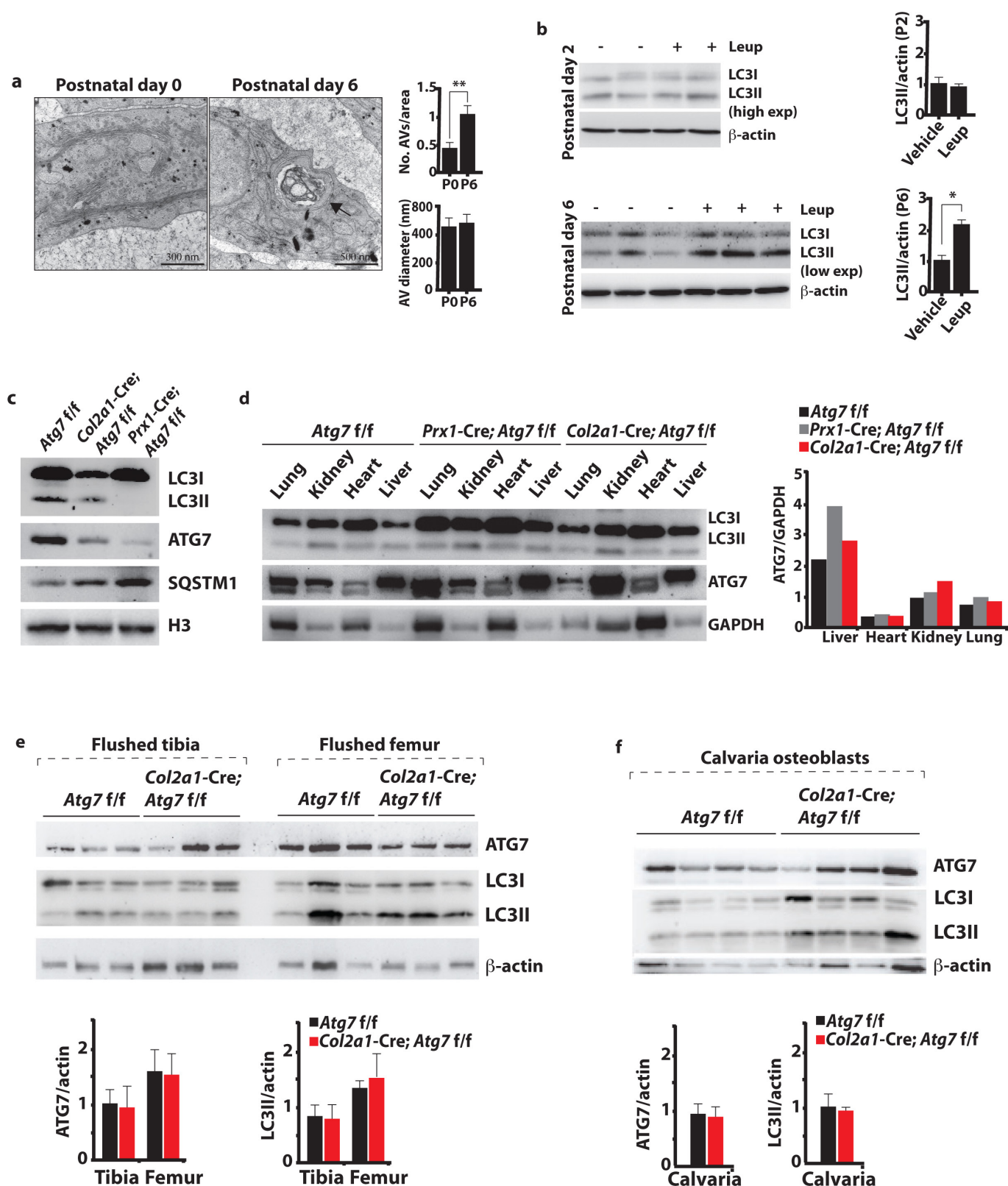
protease inhibitor tablet per 10 ml). Cell lysates were rotated at 4°C for at least 30 min, and then the soluble fraction was isolated by centrifugation at 14,000 r.p.m. for 10 min at 4°C. A fraction of the clarified lysate was used for western blot analysis. Primary antibody or rabbit pre-immune IgG were added to the lysates and rotated overnight at 4°C, and then 25 μl of Protein A Sepharose beads (Sigma-Aldrich) were added and rotated for 2 h at 4°C. Immunoprecipitates were washed three times with cold lysis buffer. Whole-cell lysates and immunoprecipitated proteins were boiled in 30 μl sample buffer, separated by SDS-PAGE on precast 4–15% gels (BioRad), transferred on PVDF membranes and probed with antibodies against beclin-1 (Santa Cruz Biotechnology), VPS34 (Sigma-Aldrich) and BCL2 (Cell Signaling Technology), FGFR3 (Santa Cruz), FGFR4 (Cell Signaling) and p-tyrosine (4G10, Millipore).

**PtdIns(3)K assay.** PtdIns(3)K activity in the beclin-1 immunoprecipitates was determined using the Class III PI3K ELISA kit (Echelon Biosciences) according to the manufacturer's instructions. Immunocomplexes were incubated with a reaction mixture containing phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5)P<sub>2</sub>) substrate and ATP for 3 h, and the amount of PtdIns(3,4,5)P<sub>3</sub> generated from PtdIns(4,5)P<sub>2</sub> by PtdIns(3)K was quantified using a competitive ELISA. Equal amounts of beclin-1 immunoprecipitate were evaluated by western blotting using beclin-1 antibody.

**Procollagen secretion assay.** To follow PC2 secretion in RCS chondrocytes, cells were pre-treated overnight with ascorbic acid (100 μg ml<sup>-1</sup>) in DMEM without FCS. Cells were then labelled with 37.5 μCi ml<sup>-1</sup> 2,3-<sup>3</sup>H-Proline (Perkin Elmer) for 4 h at 40°C in DMEM, without FCS supplemented with ascorbic acid (100 μg ml<sup>-1</sup>), then shifted to 32°C in DMEM without FCS containing cold proline (10 mM), 20 mM HEPES pH 7.2 and ascorbic acid (100 μg ml<sup>-1</sup>). After 0, 30 and 60 min the medium and cells were collected, lysed and proteins precipitated in saturated ammonium sulfate overnight, then resuspended in Laemmli buffer. Samples were run on 4–15% precast gels (Biorad), transferred onto nitrocellulose membrane (Whatman, Perkin Elmer) and developed by autoradiography using the BetaMAGER-D system and analysed using M3 Vision software (Biospace Lab).

**Statistics.** Paired Student's *t*-test was performed when comparing the same cell population with two different treatments. Unpaired Student's *t*-test was performed when comparing two groups of mice or different primary chondrocyte preparations. One-way ANOVA and Tukey's post-hoc tests were performed when comparing more than two groups relative to a single factor (time or treatment/genotype). Two-way ANOVA and Tukey's post-hoc tests were performed when comparing more than two groups relative to two factors (time and treatment/genotype). No statistical methods were used to predetermine sample size.

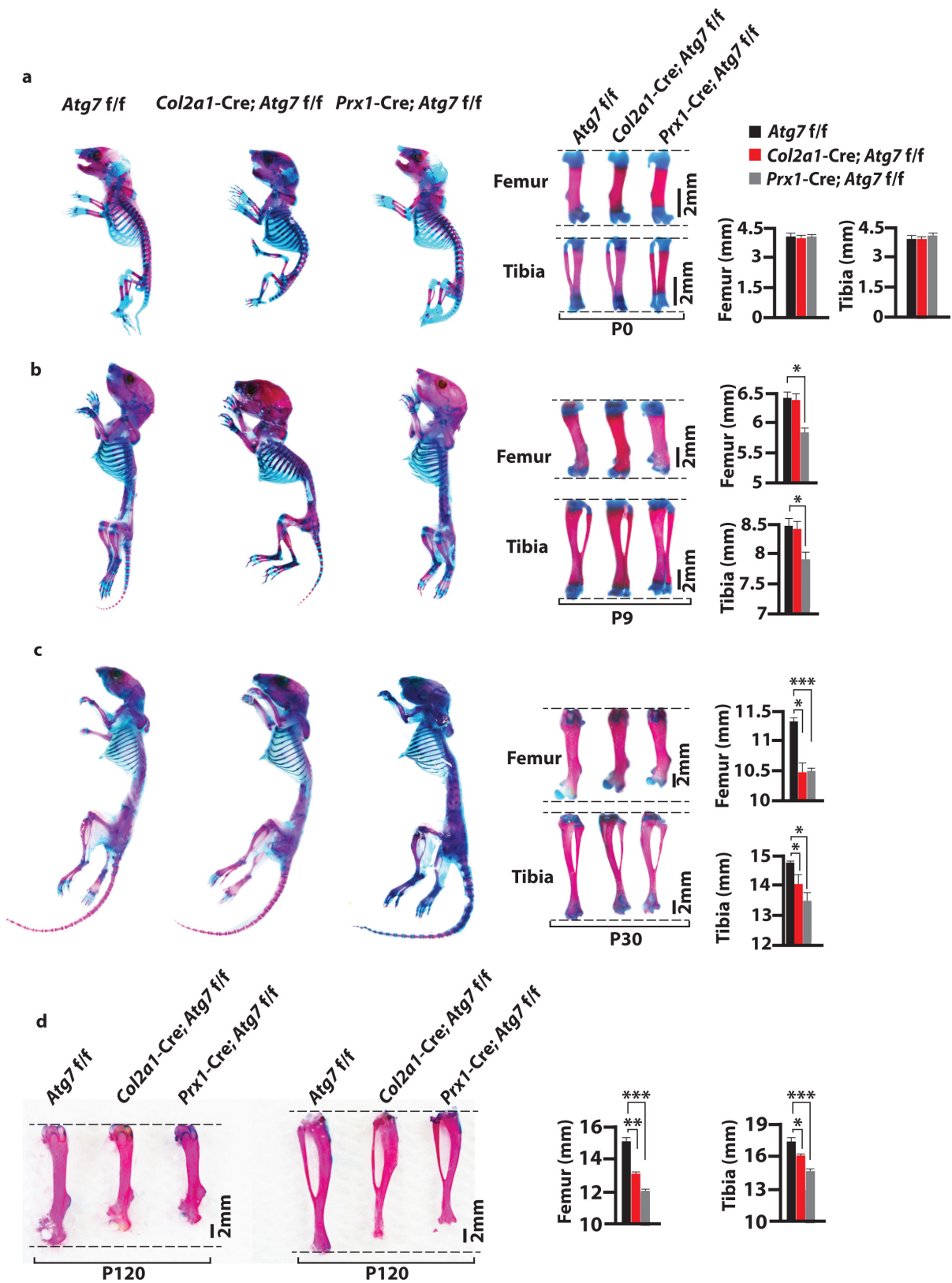
29. Venditti, R. *et al.* Sedlin controls the ER export of procollagen by regulating the Sar1 cycle. *Science* **337**, 1668–1672 (2012).
30. King, K. B. & Kimura, J. H. The establishment and characterization of an immortal cell line with a stable chondrocytic phenotype. *J. Cell. Biochem.* **89**, 992–1004 (2003).



**Extended Data Figure 1 | Autophagy flux increases during early post-natal bone development.** **a**, Representative images of comparative TEM of P0 and P6 wild-type growth-plate chondrocytes showing increased AV number at P6. Arrow indicates AVs. Bar graphs show number and size of AVs within 5.3  $\mu$ m field of view. Values represent mean  $\pm$  s.e.m. Student's *t*-test,  $**P < 0.005$ . *N* = 16 (P0); 18 (P6).

**b**, Western blot analysis of LC3I/II of femoral growth plates from mice at the indicated ages. Mice were injected with Leupeptin where indicated (Leup; 40 mg kg<sup>-1</sup>; 6 h before being killed).  $\beta$ -Actin was used as a loading control. Bar graphs show quantification of LC3II protein

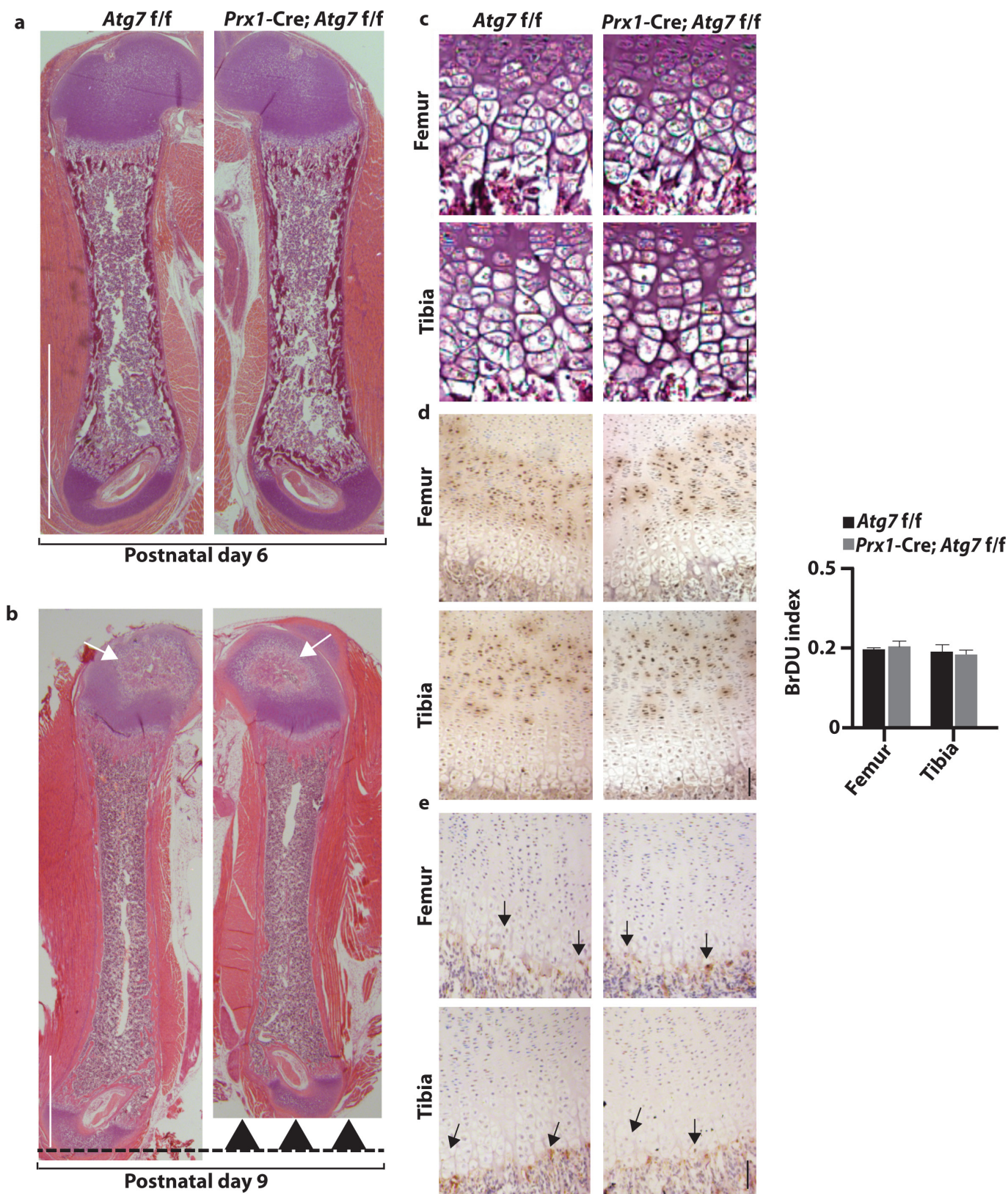
relative to  $\beta$ -actin. Mean  $\pm$  s.e.m. Student's *t*-test,  $*P < 0.05$ . *n* = 3 per group. **c**, Western blot analysis of ATG7, LC3 and SQSTM1 proteins in Atg7<sup>f/f</sup>, Col2a1-Cre; Atg7<sup>f/f</sup> and Prx1-Cre; Atg7<sup>f/f</sup> growth-plate lysates. Histone 3 (H3) was used as loading control. Results are representative of 3 independent experiments. **d–f**, Western blot analysis of ATG7 and LC3 proteins isolated from different tissues isolated from mice with indicated genotypes. GAPDH and  $\beta$ -actin were used as loading control. Bar graph shows quantification ( $\pm$  s.e.m.) of ATG7 and LC3II proteins relative to  $\beta$ -actin in different tissues. *n* = 2 (**d**), 3 (**e**), 4 (**f**) mice per genotype.



**Extended Data Figure 2 | Analysis of bone growth in mice lacking chondrocyte autophagy.** a–d, Alcian blue/alizarin red skeletal staining of Atg7<sup>f/f</sup>, Col2a1-Cre; Atg7<sup>f/f</sup> and Prx1-Cre; Atg7<sup>f/f</sup> mice at P0 (a), P9 (b), P30 (c) and P120 (d). Details of femur and tibia magnifications. In c and d

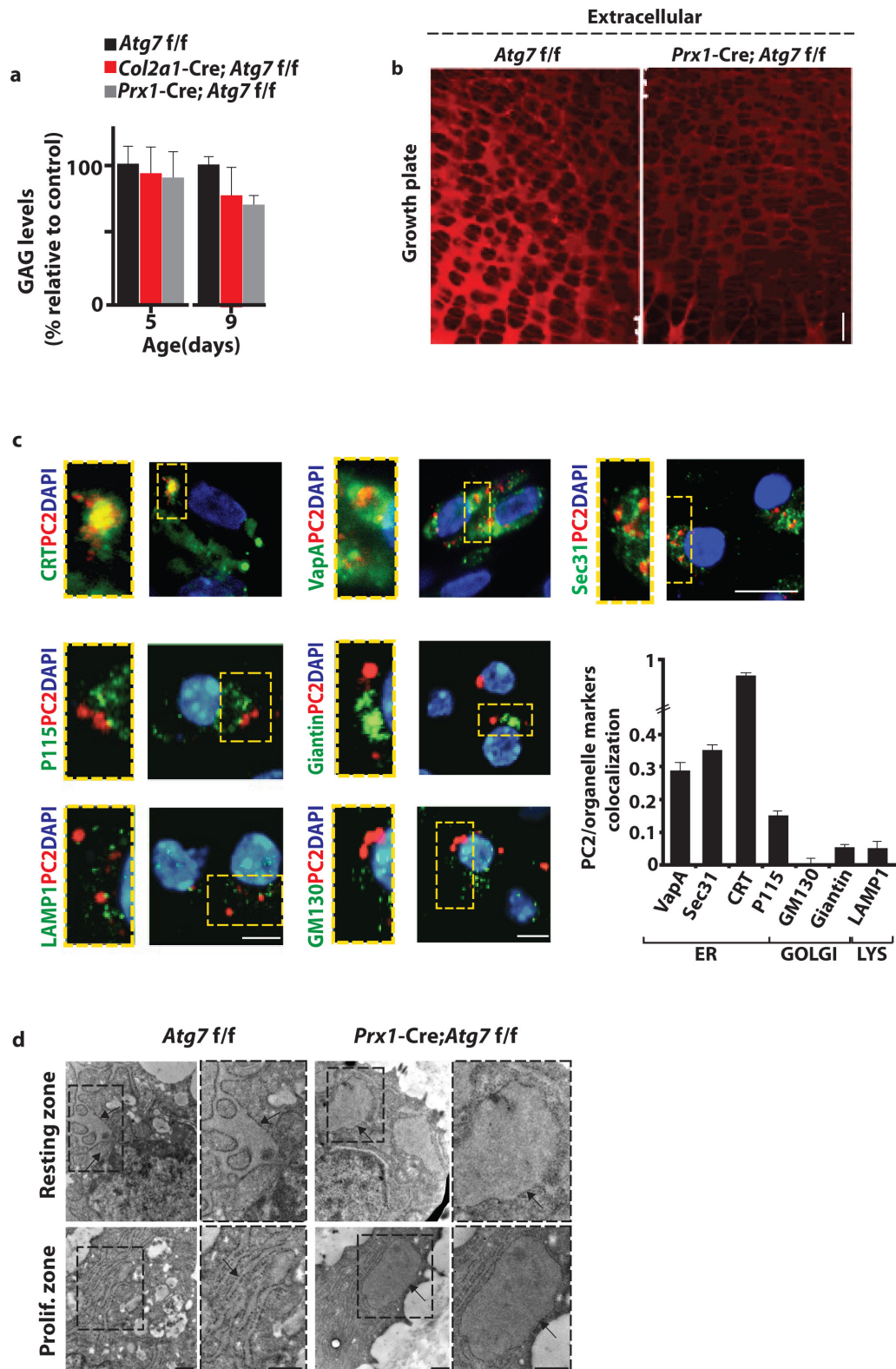
only male mice were analysed. Graphs show femur and tibia lengths from mice with indicated genotypes. Values represent mean  $\pm$  s.e.m. Student's *t*-test, \* $P < 0.05$ , \*\* $P < 0.005$ , \*\*\* $P < 0.0005$ .  $n = 3$  (P0), 3 (P9), 4 (P30), 4 (P120) mice per genotype. Scale bar, 2 mm.





**Extended Data Figure 3 | Autophagy does not regulate chondrocyte proliferation, differentiation and apoptosis.** **a, b,** Haematoxylin and eosin (H&E) staining of femoral sections of P6 (**a**) and P9 (**b**) *Atg7<sup>fl/fl</sup>* and *Prx1-Cre; Atg7<sup>fl/fl</sup>* mice showing a reduced femoral length starting from P9 in *Prx1-Cre; Atg7<sup>fl/fl</sup>* mice compared with controls (black arrowheads). White arrows show normal differentiation of the secondary ossification centre in *Prx1-Cre; Atg7<sup>fl/fl</sup>* compared with control. The data are representative of three independent experiments. Scale bar, 2 mm.

**c–e,** Representative images of H&E staining of hypertrophic chondrocytes (**c**), bromodeoxyuridine (BrdU) staining (**d**) and TUNEL assay (**e**) in P6 *Atg7<sup>fl/fl</sup>* and *Prx1-Cre; Atg7<sup>fl/fl</sup>* growth plates. Arrows indicate TUNEL-positive cells. Graph shows quantification of BrdU index in femoral and tibial growth plates from *Atg7<sup>fl/fl</sup>* and *Prx1-Cre; Atg7<sup>fl/fl</sup>* mice. Values represent  $\pm$  s.e.m.  $n = 5$  mice per genotype. Scale bar, 100  $\mu$ m. **c, e,** Data are representative of 3 independent experiments.

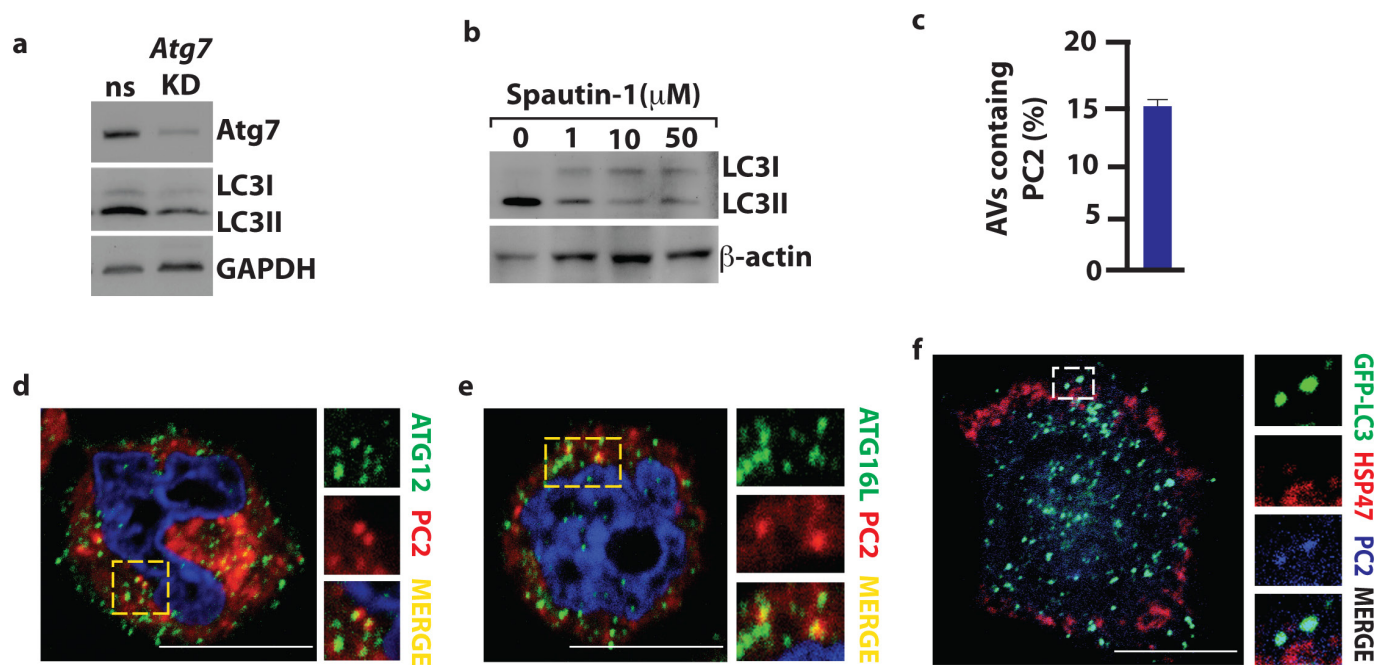


#### Extended Data Figure 4 | Lack of autophagy leads to PC2

**accumulation in the ER of growth-plate chondrocytes.** **a**, Total levels of glycosaminoglycans (GAGs) in femoral and tibia growth plates of P5 and P9 mice with indicated genotypes. Values were normalized to total DNA and expressed as percentage relative to control (*Atg7*<sup>f/f</sup>) mice. Values represent mean  $\pm$  s.e.m.  $n = 3$  (P5);  $n = 5$  (P9) mice per genotype. **b**, Extracellular Col2 staining in chondroitinase ABC-treated growth-plate femoral sections isolated from *Atg7*<sup>f/f</sup> and *Prx1-Cre; Atg7*<sup>f/f</sup> mice. The data are representative of two independent experiments. Scale bar, 500  $\mu$ m. **c**, Representative images of confocal analysis of PC2, calreticulin (CRT)

Sec31, VapA, P115, GM130, giantin and LAMP1 proteins in *Prx1-Cre; Atg7*<sup>f/f</sup> growth-plate chondrocytes at P6. Insets show high magnification of boxed areas. Scale bar, 10  $\mu$ m. Quantification ( $\pm$  s.e.m.) of PC2 co-localization with indicated organelle markers (Mander's coefficient, ImageJ plug-in). At least 2 sections per mouse containing 400 cells were analysed.  $n = 3$  mice per group. **d**, ER distention in *Atg7*<sup>f/f</sup>; *Prx1-Cre* growth-plate chondrocytes compared with control chondrocytes (*Atg7*<sup>f/f</sup>). Arrows indicate ER. Insets show a high magnification of selected areas. The data are representative of two independent experiments. Scale bar, 500 nm.



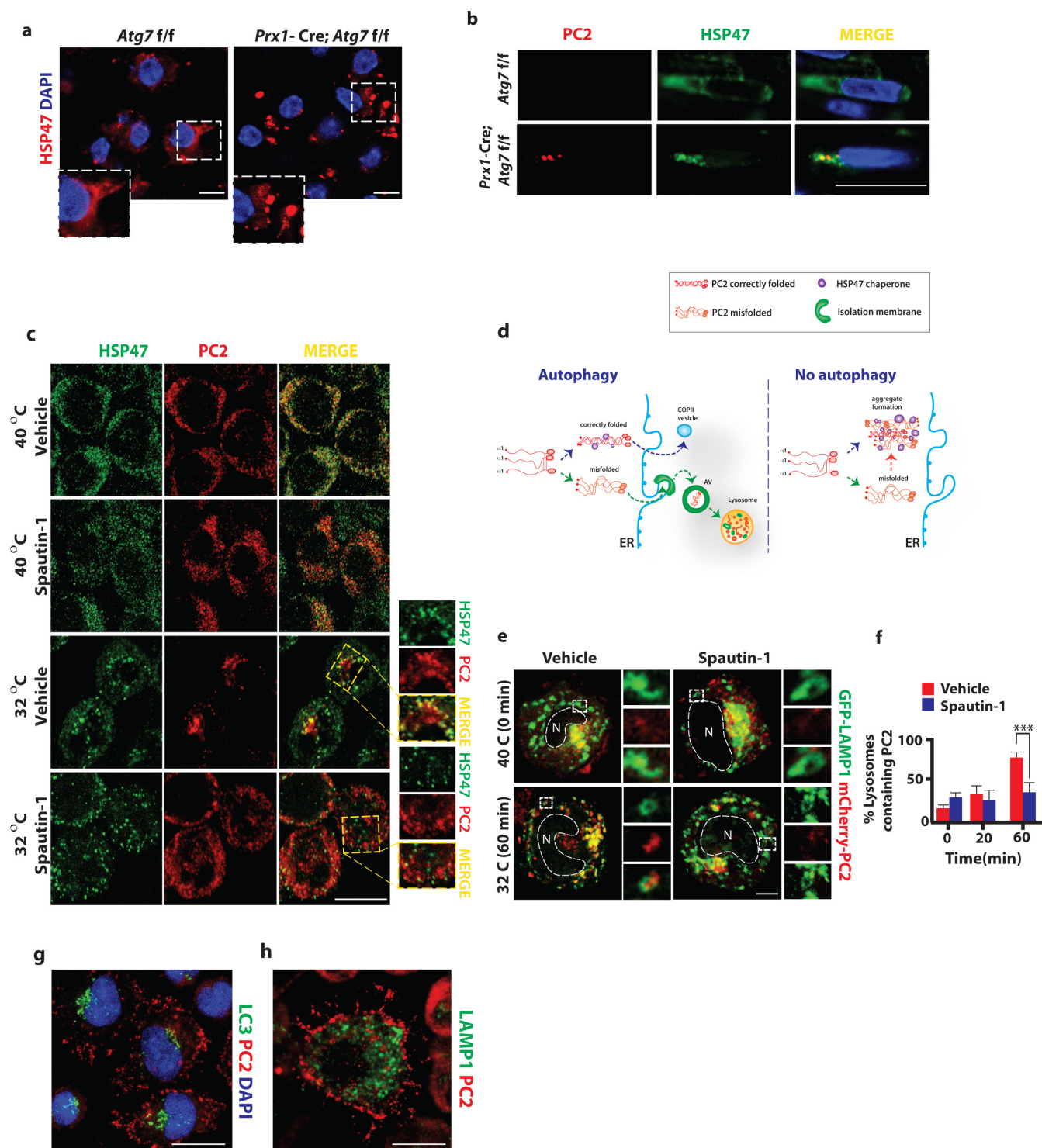


#### Extended Data Figure 5 | Autophagy regulates PC2 secretion.

**a**, Western blot analysis of ATG7 and LC3II levels in control (scrambled) and *Atg7* siRNA-treated RCS chondrocytes. GAPDH was used as a loading control. KD, knockdown; ns, non-silencing. **b**, Western blot analysis of LC3II levels in RCS chondrocytes treated with Spautin-1 at the indicated concentrations for 24 h.  $\beta$ -Actin was used as a loading control. **c**, Co-localization of GFP-LC3 with PC2 in primary chondrocytes isolated from GFP-LC3 transgenic mice and treated with BafA1 (4 h, 200 nM). Values are expressed as percentage  $\pm$  s.d. relative to total GFP area.

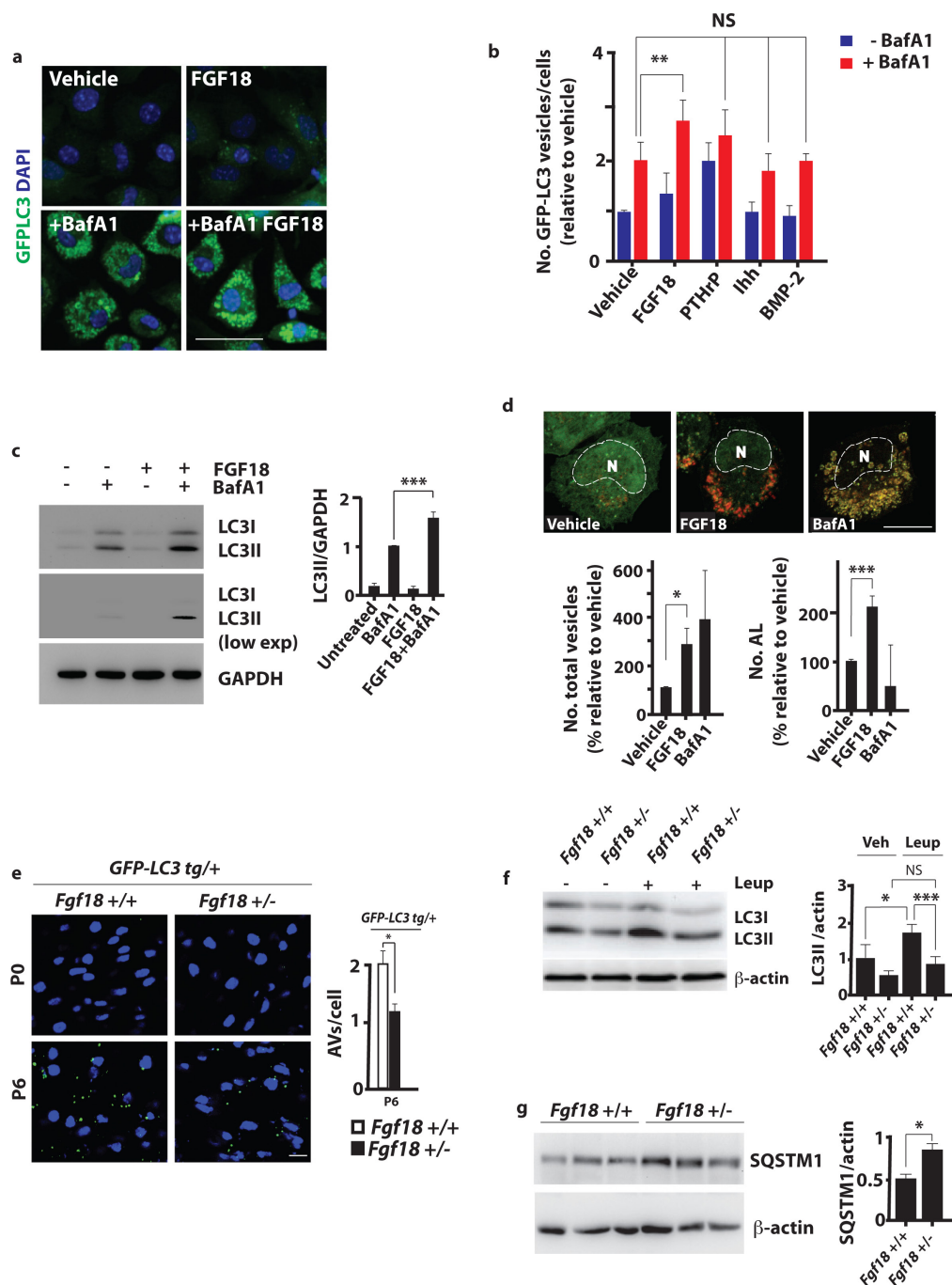
*N* = 500 cells; *n* = 2 independent preparations. **d, e**, Confocal analysis of ATG12 (d), ATG16L (e) (green) and mCherry-PC2 (red) in RCS chondrocytes. The data are representative of 3 independent experiments. Insets show a high magnification of selected areas. Blue, DAPI. Scale bar, 10  $\mu$ m. **f**, Immunofluorescence staining of PC2 (blue), HSP47 (red) and GFP-LC3 (green) in RCS chondrocytes, showing that HSP47 does not co-localize with PC2 in AVs. The data are representative of 2 independent experiments. Insets show higher magnifications and single colour channels of the boxed area. Scale bar, 5  $\mu$ m.





**Extended Data Figure 6 | Role of autophagy during PC2 secretion.**  
**a**, Immunofluorescence staining of HSP47 chaperone (red) in *Atg7<sup>fl/fl</sup>* and *Atg7<sup>fl/fl</sup>; Prx1-Cre* chondrocytes, showing altered HSP47 distribution in *Prx1-Cre; Atg7<sup>fl/fl</sup>* chondrocytes. The data are representative of 3 independent experiments. Insets show a higher magnification of the boxed area. Scale bar, 10  $\mu$ m. Blue, DAPI. **b**, Co-localization of PC2 with HSP47 in growth-plate chondrocytes of mice with indicated genotypes. The data are representative of 3 independent experiments. Scale bar, 20  $\mu$ m. **c**, Altered HSP47 and PC2 trafficking in Spautin-1-treated chondrocytes. HSP47 and PC2 immunostaining in control (vehicle) or Spautin-1-treated RCS chondrocytes. Synchronized PC2 secretion was obtained after incubating chondrocytes for 3 h at 40 °C to block PC2 in the ER, and then shifting the temperature to 32 °C (ER block release) for 10 min. The data are representative of 2 independent experiments. Scale bar, 10  $\mu$ m.

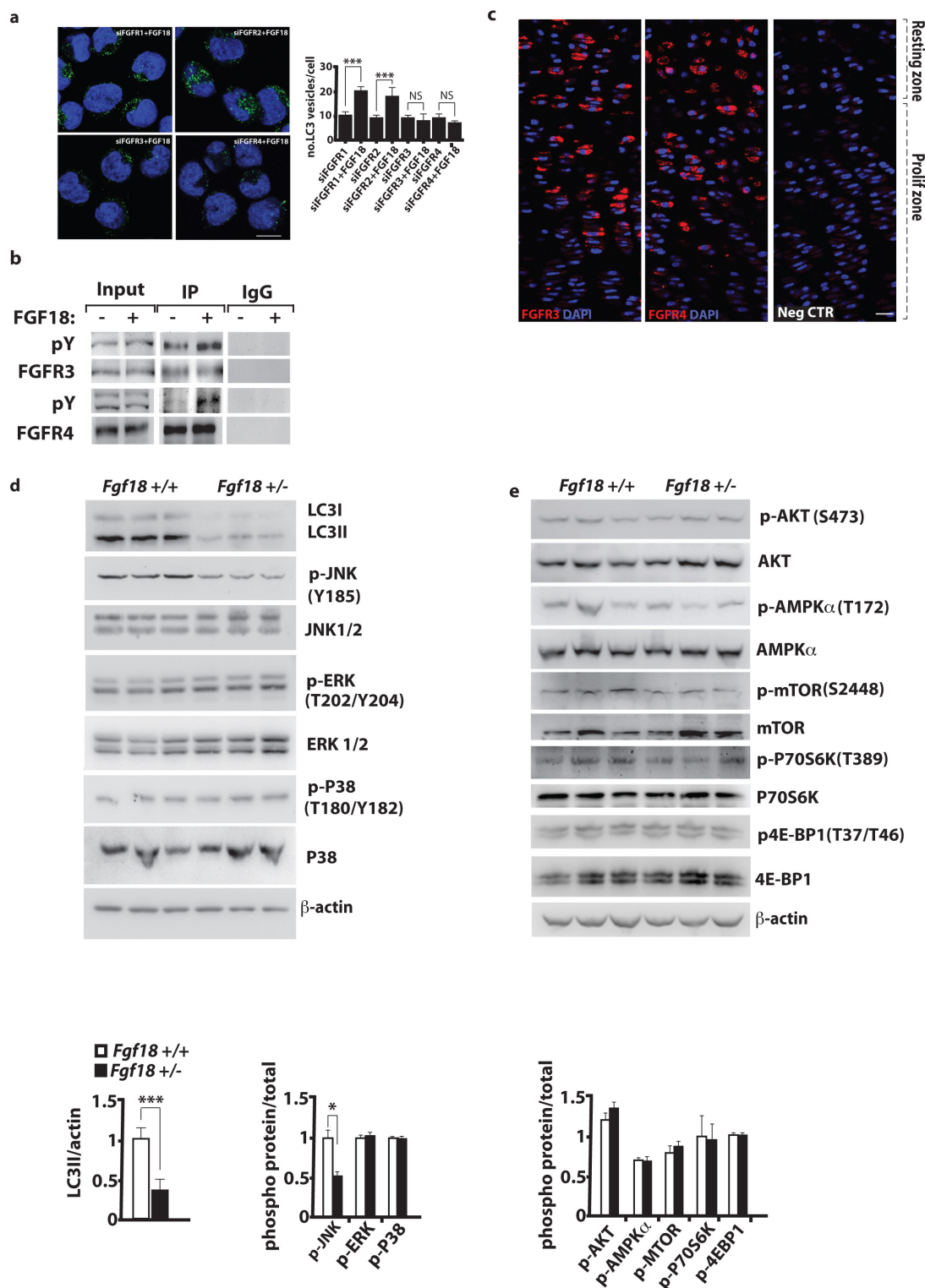
**d**, Proposed model of autophagy function in chondrocytes. Autophagy in chondrocytes prevents PC2 aggregation and maintains ER homeostasis during the process of PC2 secretion. **e**, Confocal analysis of GFP-LAMP1 (green) and mCherry-PC2 (red) in vehicle- and Spautin-1-treated chondrocytes at the indicated time points (min) after the ER block release. The insets show a high magnification of the selected area. Scale bar, 5  $\mu$ m. **f**, Quantification of GFP-LAMP1/mCherry-PC2 co-localization. Values represent mean  $\pm$  s.d. from three independent experiments.  $N = 30$ . ANOVA,  $P = 4.91 \times 10^{-5}$ ; Tukey's post-hoc test, \*\*\* $P < 0.0005$ . **g, h**, Confocal analysis of RCS chondrocytes treated with tannic acid (0.5% final concentration in the medium) for 1 h, showing that PC2 vesicles (red) at the periphery do not co-localize with LC3 (**g**) or with LAMP1 (**h**) (green). The data are representative of 2 independent experiments. Scale bar, 10  $\mu$ m.



### Extended Data Figure 7 | FGF18 induces autophagy in chondrocytes.

**a**, Representative images of high-content imaging analysis of primary chondrocytes isolated from GFP-LC3 transgenic mice and treated with vehicle or FGF18 (25 ng ml<sup>-1</sup> for 24 h). BafA1 was used where indicated (4 h, 200 nM). Scale bar, 50 μm. **b**, Quantification of green vesicles (AVs) in cells treated with the indicated factors for 24 h. Vesicles were counted in at least 1,000 cells per treatment. Values represent mean values ± s.d. of *n* = 3 independent experiments. Statistical analysis was performed using repeated-measures ANOVA, *P* = 0.0001; Tukey's post-hoc test, \*\**P* < 0.005. **c**, Western blot analysis of primary chondrocytes isolated from wild-type mice treated as indicated (FGF18, 25 ng ml<sup>-1</sup>, 24 h). Where indicated, BafA1 was added (200 nM, 4 h). Bar graphs represent mean values ± s.e.m. of *n* = 3 independent experiments. Student's *t*-test, \*\*\**P* < 0.0005. **d**, Representative images of immunofluorescence staining of RCS chondrocytes expressing the tandem fluorescent-tagged LC3 (mRFP-eGFP-LC3) protein. Graphs show increased number of autolysosomes (AL) and of total vesicles (AV+AL) in FGF18- (25 ng ml<sup>-1</sup> for 24 h) compared with vehicle-treated RCS chondrocytes. Bar graph represent mean values ± s.d. *N* = 10 cells per experiment were

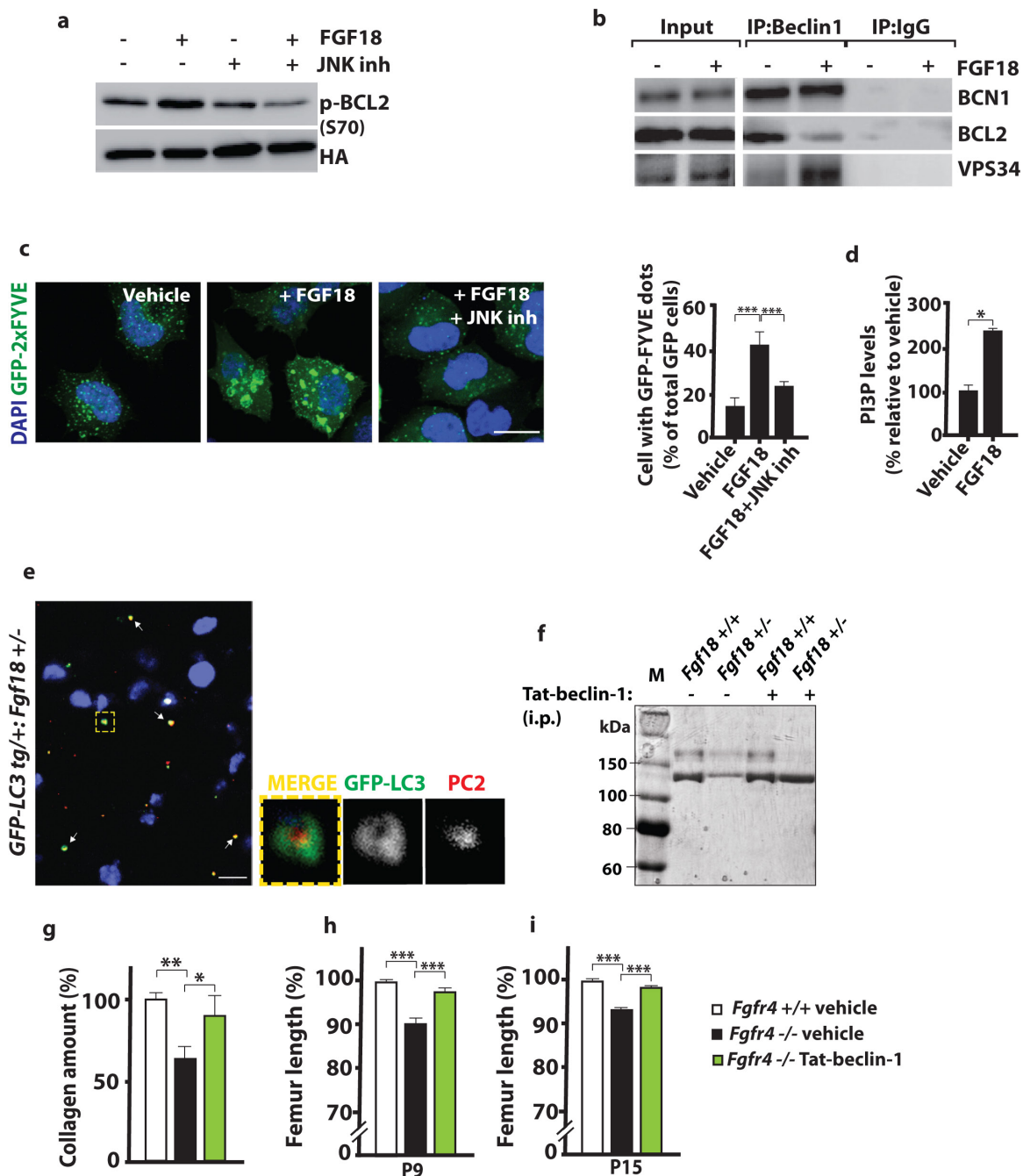
analysed from 3 independent experiments. Student's *t*-test, \**P* < 0.05, \*\*\**P* < 0.0005. BafA1 (4 h at 200 nM) was used as a control to inhibit AV-lysosome fusion. Scale bar, 10 μm. **e**, Representative images of confocal analysis of GFP-LC3 puncta (autophagosomes) in femoral growth plates from P0 and P6 GFP-LC3<sup>tg/+</sup>; Fgf18<sup>+/+</sup> and GFP-LC3<sup>tg/+</sup>; Fgf18<sup>+/-</sup> mice. Scale bar, 20 μm. Bar graph shows quantification of the data in P6 mice. Values are mean ± s.e.m. *n* = 5 mice per group. Two sections per mouse containing at least 400 nuclei were analysed. Student's *t*-test, \**P* < 0.05. **f**, Western blot analysis of Fgf18<sup>+/+</sup> and Fgf18<sup>+/-</sup> growth plate lysates. Mice were injected with leupeptin where indicated (Leup; 40 mg kg<sup>-1</sup>, 6 h before being killed). β-Actin was used as loading control. Bar graph shows quantification of LC3II protein in vehicle and leupeptin-injected mice. Values represent the mean values (± s.e.m.) relative to β-actin. *n* = 5 mice per genotype. ANOVA, *P* = 7.17 × 10<sup>-6</sup>; Tukey's post-hoc test, \**P* < 0.05, \*\*\**P* < 0.0005. **g**, Western blot analysis of SQSTM1 protein in three Fgf18<sup>+/+</sup> and three Fgf18<sup>+/-</sup> growth plate lysates at P30. β-Actin was used as a loading control. Bar graph shows quantification of SQSTM1 protein relative to β-actin. Values represent mean ± s.e.m. *n* = 3 mice per group. Student's *t*-test, \**P* < 0.05.



**Extended Data Figure 8 | FGF18 regulates autophagy via FGFR4 and JNK1/2 signalling.** **a**, Representative images of immunofluorescence analysis of LC3 positive vesicles in RCS chondrocytes treated with siRNA for *Fgfr1*, *Fgfr2*, *Fgfr3* and *Fgfr4* and then stimulated with FGF18 for 2 h. BafA1 was added (200 nM, 3 h). Values represent mean values  $\pm$  s.e.m. of  $n = 3$  independent experiments ( $N = 40$  cells per treatment were analysed). Student's  $t$ -test, \*\*\* $P < 0.0005$ . NS, not significant. Scale bar, 10  $\mu$ m. **b**, Immunoprecipitation of FGFR3 or of FGFR4 from RCS chondrocytes stably expressing FGFR3 or FGFR4, respectively, followed by western blotting with phosphotyrosine antibody (pY). Cells were untreated (–) or treated (+) with FGF18 (100 ng ml<sup>-1</sup>, 20 min). **c**, Confocal analysis of FGFR3 and FGFR4 in growth-plate chondrocytes isolated from P6 mice. No signal was detected when sections were incubated with

secondary antibody alone (Neg. CTR). The data are representative of two independent experiments. Scale bar, 20  $\mu$ m. **d**, Western blot analysis of LC3I/II, phospho-JNK1/2, JNK1/2, phospho-ERK1/2, ERK1/2, phospho-P38 MAPK and P38 MAPK in growth plates isolated from three *Fgf18*<sup>+/+</sup> and three *Fgf18*<sup>+/-</sup> mice at P6.  $\beta$ -Actin was used as a loading control. The bar graph shows quantification of LC3II relative to  $\beta$ -actin and of phosphorylated proteins relative to the corresponding total proteins. Values are mean  $\pm$  s.e.m. from  $n = 3$  mice per genotype. Student's  $t$ -test, \* $P < 0.05$ , \*\*\* $P < 0.0005$ . **e**, Western blot analysis of three *Fgf18*<sup>+/+</sup> and three *Fgf18*<sup>+/-</sup> growth-plate lysates showing no differences in the phosphorylation levels of the proteins analysed. Bar graph shows quantification of the ratio of phosphorylated to total protein (values represent mean  $\pm$  s.e.m.;  $n = 3$ ).

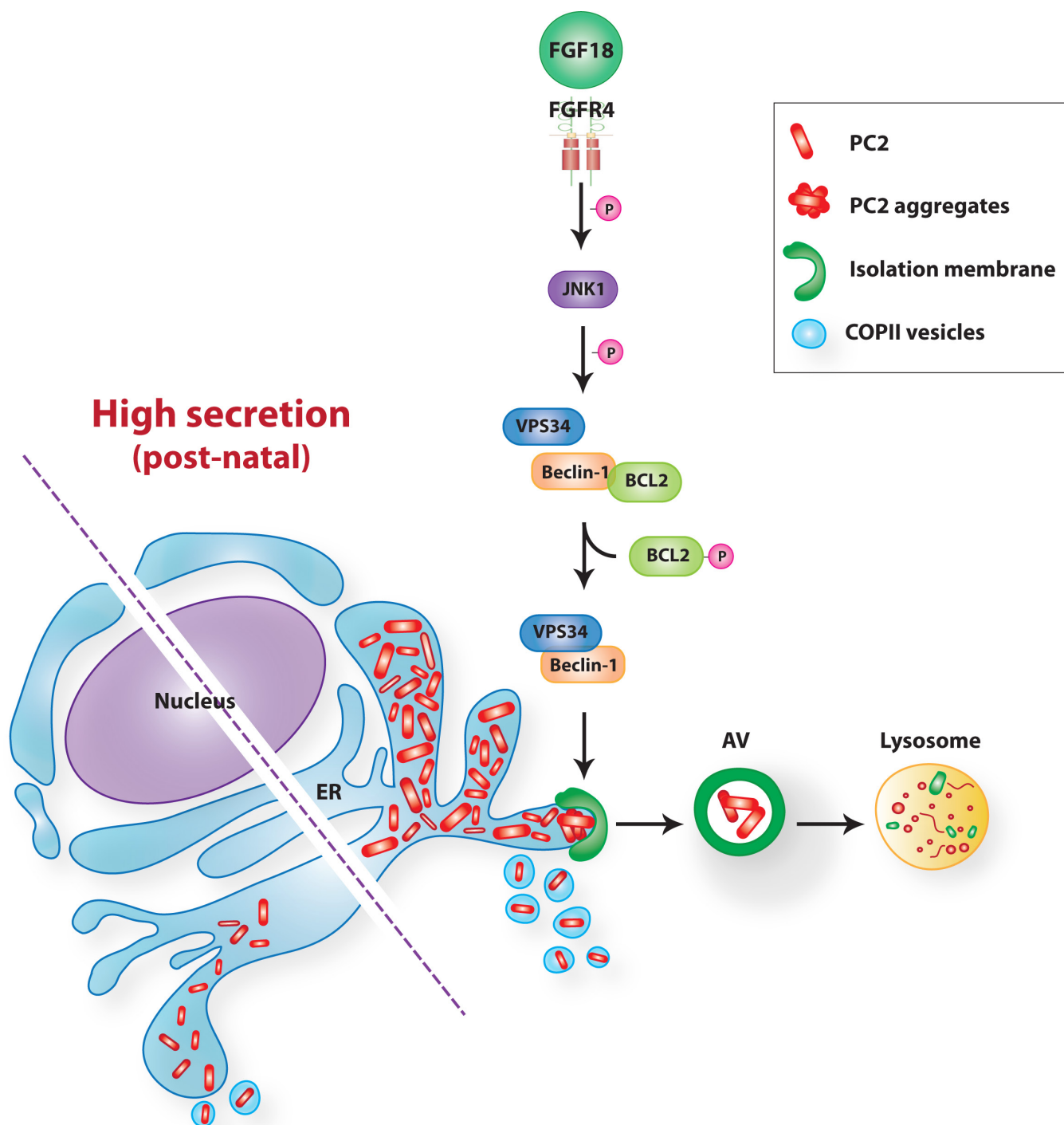




### Extended Data Figure 9 | FGF18 regulates autophagy through beclin-1 complex activation.

**a**, Western blot analysis of BCL2 phosphorylation at Ser 70, p-BCL2 (S70) and of human influenza hemagglutinin (HA) in RCS chondrocytes expressing human BCL2-HA. Where indicated, chondrocytes were treated with FGF18 (2 h, 25 ng ml<sup>-1</sup>) and with JNK inhibitor (JNK inh; 4 h, 50 μM). **b**, Immunoprecipitation assays testing physical interactions between endogenous beclin-1, BCL2 and VPS34 in untreated and FGF18-treated (2 h, 25 ng ml<sup>-1</sup>) RCS chondrocytes. Cell lysates were immunoprecipitated with a beclin-1-specific antibody or control immunoglobulin G (IgG), followed by probing with antibodies specific for beclin-1, BCL2 or VPS34. **c**, Representative images of membrane-associated PtdIns(3)K assay *in situ*. RCS chondrocytes were transfected with GFP-2•FYVE and then treated with or without FGF18 (2 h, 25 ng ml<sup>-1</sup>) and with JNK inhibitors (4 h, 50 μM) where indicated. Graph shows quantitative analysis of cells with GFP-2•FYVE dots. Values are expressed as mean (± s.e.m.) of *n* = 3 independent experiments. *N* = 65 (vehicle), 237 (FGF18), 203 (FGF18 + JNK inhibitor) cells. ANOVA, *P* = 1.12 × 10<sup>-6</sup>; Tukey's post-hoc test, \*\*\**P* < 0.0005. Scale bar, 10 μm. **d**, Measure of PtdIns(3)K activity associated with beclin-1, expressed

as fold change relative to control cells (vehicle treated). Graph shows mean ± s.e.m. Student's *t*-test, \**P* < 0.05. **e**, PC2 (red) and GFP-LC3 (green) confocal analysis of resting chondrocytes in P6 GFP-LC3<sup>tg/+</sup>; Fgf18<sup>+/-</sup> mice showing autophagosomes containing PC2 (arrows). The data are representative of 2 independent experiments. The inset shows a high magnification and single channels of the boxed areas. Scale bar, 10 μm. **f**, Coomassie blue staining of femoral growth-plate collagen isolated from Fgf18<sup>+/+</sup>, Fgf18<sup>+/-</sup> and Fgf18<sup>-/-</sup> mice injected with Tat-beclin-1. M, marker. **g**, Total collagen concentration in femoral and tibia growth plates of Fgfr4<sup>+/+</sup>, Fgfr4<sup>-/-</sup> and Fgfr4<sup>-/-</sup> mice treated with Tat-beclin-1 at P9. Values (mean ± s.e.m.) were normalized to DNA and expressed as percentage relative to control mice. *n* = 5 mice per group. ANOVA, *P* = 0.0004; Tukey's post-hoc test, \*\**P* < 0.005, \*\*\**P* < 0.0005. **h**, **i**, Femoral lengths of Fgfr4<sup>+/+</sup>, Fgfr4<sup>-/-</sup> and Fgfr4<sup>-/-</sup> mice treated with Tat-beclin-1 at P9 (**h**) and P15 (**i**). Values (mean ± s.e.m.) were expressed as percentage relative to littermate control mice. *n* = 5 mice per group (**h**) and *n* = 4 mice per group (**i**). ANOVA, *P* = 8.8 × 10<sup>-10</sup> (**h**), *P* = 5.83 × 10<sup>-5</sup> (**i**); Tukey's post-hoc test, \*\*\**P* < 0.0005.



**Extended Data Figure 10 | Proposed model of FGF18-dependent regulation of autophagy in chondrocytes during post-natal bone growth.** During early post-natal bone growth, FGF18 induces the activation of FGFR4 and of JNK kinase, which phosphorylates BCL2 and activates the

VPS34-beclin-1 autophagy complex. This process induces autophagy, which maintains PC2 homeostasis by preventing accumulation of PC2 in the ER. Chondrocyte autophagy appears to be dispensable when low levels of PC2 secretion are needed (for example, during prenatal bone growth).

# Force generation by skeletal muscle is controlled by mechanosensing in myosin filaments

Marco Linari<sup>1,2</sup>, Elisabetta Brunello<sup>1,†</sup>, Massimo Reconditi<sup>1,2</sup>, Luca Fusi<sup>3</sup>, Marco Caremani<sup>1</sup>, Theyencheri Narayanan<sup>4</sup>, Gabriella Piazzesi<sup>1</sup>, Vincenzo Lombardi<sup>1</sup> & Malcolm Irving<sup>3</sup>

**Contraction of both skeletal muscle and the heart is thought to be controlled by a calcium-dependent structural change in the actin-containing thin filaments, which permits the binding of myosin motors from the neighbouring thick filaments to drive filament sliding<sup>1–3</sup>. Here we show by synchrotron small-angle X-ray diffraction of frog (*Rana temporaria*) single skeletal muscle cells that, although the well-known thin-filament mechanism is sufficient for regulation of muscle shortening against low load, force generation against high load requires a second permissive step linked to a change in the structure of the thick filament. The resting (switched ‘OFF’) structure of the thick filament is characterized by helical tracks of myosin motors on the filament surface and a short backbone periodicity<sup>2,4,5</sup>. This OFF structure is almost completely preserved during low-load shortening, which is driven by a small fraction of constitutively active (switched ‘ON’) myosin motors outside thick-filament control. At higher load, these motors generate sufficient thick-filament stress to trigger the transition to its long-periodicity ON structure, unlocking the major population of motors required for high-load contraction. This concept of the thick filament as a regulatory mechanosensor provides a novel explanation for the dynamic and energetic properties of skeletal muscle. A similar mechanism probably operates in the heart.**

Muscle contraction is driven by the relative sliding of two sets of overlapping filaments, the myosin-containing thick filaments and the actin-containing thin filaments, coupled to ATP hydrolysis. In the widely accepted model of skeletal muscle regulation, excitation of a motor nerve causes release of calcium ions within the muscle cell that bind to troponin in the thin filaments, triggering a structural change that allows binding of the myosin motor domains of the thick filaments<sup>1</sup>. In resting muscle, however, the myosin motors are packed into helical tracks on the surface of the thick filaments<sup>2–6</sup>, making them unavailable for actin binding or ATP hydrolysis<sup>7</sup>, implying the presence of a second thick-filament-based switch for contraction. The fact that the length and axial periodicity of the thick filaments increase slightly on muscle activation<sup>2,4,8,9</sup> or when resting muscle is stretched<sup>4,8–10</sup> suggests that the thick-filament switch might be mechanical, and that unpacking of the myosin motors on muscle activation might be triggered by thick-filament stress.

We used synchrotron small-angle X-ray diffraction of isolated muscle cells to test this idea (Fig. 1, see Methods). When an isolated muscle cell is activated at constant length (Fig. 1a, thick lines), force and filament stress start to increase about 10 ms after the start of stimulation (the latent period), when activation of the troponin complex in the thin filaments is already maximal<sup>3,11,12</sup>. To determine the effect of thick-filament stress on its structure during activation, we imposed rapid shortening of about 5% fibre length (Fig. 1a, thin

line in left section) to delay development of force and filament stress for about 20 ms.

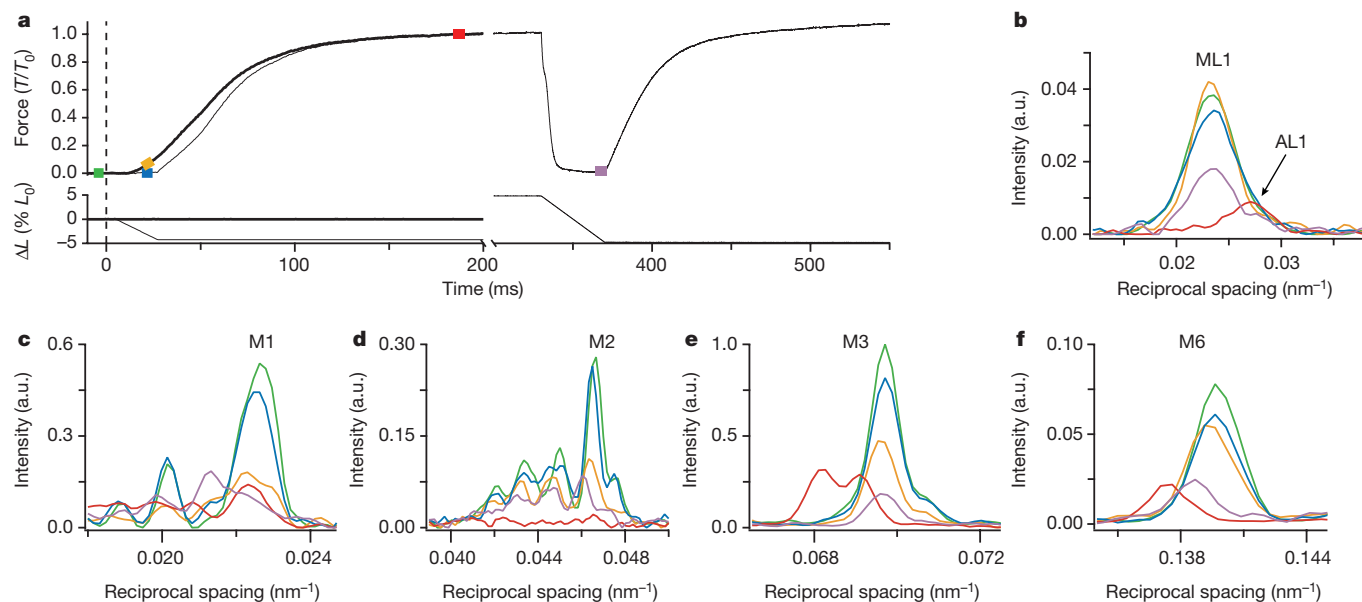
The OFF structure of the thick filament, seen in resting muscle, has a characteristic X-ray signature. The three-stranded helical arrangement of the myosin motors on the filament surface produces a series of layer-line reflections (of which the first order is called ML1) that are orders of a fundamental axial periodicity of about 43 nm, and corresponding axial reflections (of which the first order is called M1)<sup>2</sup>. Two other thick-filament components, myosin binding protein-C (MyBP-C) and titin follow the ~43-nm periodicity, and MyBP-C makes a substantial contribution to the M1 (ref. 13). The M2 reflection signals systematic perturbations in the helical arrangement. The M3, from the axial repeat of the motors, has a periodicity ( $S_{M3}$ ) of 14.34 nm in the OFF state, and the M6, dominated by the periodic mass distribution in the thick-filament backbone<sup>14,15</sup>, has a periodicity ( $S_{M6}$ ) of 7.17 nm. The axial reflections exhibit multiple peaks associated with X-ray interference between the two halves of each filament<sup>8,9,14,15</sup>.

These X-ray signals show that the thick filament remains OFF when its stress is abolished by applying rapid shortening during early activation. The intensities of the ML1, M1, M2, M3 and M6 reflections during such shortening (Fig. 1b–f, blue) are the same as those at rest (green), as are the characteristic OFF periodicities  $S_{M3}$  and  $S_{M6}$ . This behaviour is in marked contrast with the ON state at maximum isometric force ( $T_0$ ) reached after 200 ms of activation (Fig. 1a, red rectangle), in which ML1 is replaced by a much weaker actin-based reflection AL1 (Fig. 1b, red), M2 is absent (Fig. 1d), M3 moves to a periodicity of 14.57 nm (Fig. 1e) and  $S_{M6}$  increases to 7.29 nm (Fig. 1f).  $S_{M6}$  and the intensity of the M2 reflection,  $I_{M2}$ , are particularly sensitive to thick-filament stress. Thus, if the force is allowed to rise to about 0.1  $T_0$  at the same time as the zero-stress data were recorded (Fig. 1a, orange rectangle),  $I_{M2}$  and  $S_{M6}$  (Fig. 1d, f, orange) have already started to move towards their fully ON values (red)<sup>16</sup> (see also Extended Data Fig. 1). These results show that the OFF structure of the thick filament is fully preserved early in the activation process, when the intracellular calcium concentration is maximal, provided that thick-filament stress is abolished. Muscle can generate shortening at the maximum velocity with most of its myosin motors switched OFF, but a small stress (0.1  $T_0$ ) is sufficient to trigger an increase in filament periodicity and exit from the OFF conformation, consistent with the mechanosensing hypothesis.

To test this idea further, we allowed the thick filament to attain its fully ON structure at force  $T_0$ , then imposed rapid shortening of 10% fibre length to hold force at zero for about 40 ms (Fig. 1a, right section). The ML1 reflection signalling the OFF structure partially recovered (Fig. 1b, purple), and its periodicity returned completely to the OFF value. The M2 reflection, which is absent at  $T_0$ , reappeared (Fig. 1d, purple). The periodicity of the M3 reflection recovered

<sup>1</sup>Laboratory of Physiology, Department of Biology, Università di Firenze, Sesto Fiorentino, 50019 Florence, Italy. <sup>2</sup>Consorzio Nazionale Interuniversitario per le Scienze Fisiche della Materia, UdR Firenze, Sesto Fiorentino, 50019 Florence, Italy. <sup>3</sup>Randall Division and BHF Centre for Research Excellence, King's College London, London SE1 1UL, UK. <sup>4</sup>European Synchrotron Radiation Facility, BP220, F-38043 Grenoble, France. <sup>†</sup>Present address: Randall Division and BHF Centre for Research Excellence, King's College London, London SE1 1UL, UK.





**Figure 1 | X-ray reflections associated with thick-filament structure in isolated skeletal muscle fibres depend on the external load during activation.** **a**, Force and fibre length change ( $\Delta L$ , expressed as a percentage of initial fibre length (%  $L_0$ )) following stimulation starting at time zero (vertical dashed line). X-ray time exposures are denoted by coloured rectangles: green, resting; blue, shortening during early activation; red, maximum isometric force ( $T_0$ ); orange, low isometric force ( $0.1 T_0$ );

purple, shortening from the maximum isometric force. **b**, Intensity profile of the ML1 and AL1 layer-line reflections from the myosin and actin helices, respectively. **c–f**, Meridional intensity profiles for the M1, M2, M3 and M6 reflections, respectively. Colours in **b–f** relate to protocols (coloured rectangles) in **a**; X-ray data added from three (green, yellow, blue) or two (red, purple) fibres with 3 m (**b, f**) or 10 m (**c–e**) camera length. a.u., arbitrary units.

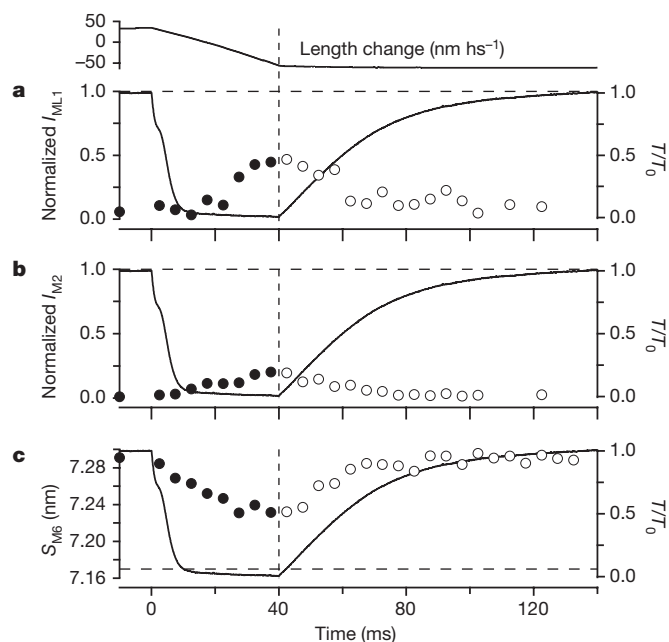
completely to its resting value of 14.34 nm (Fig. 1e, purple), and the double-peak interference effect characteristic of full activation was lost. The filament backbone periodicity ( $S_{M6}$ ; Fig. 1f) partially recovered from 7.29 nm at  $T_0$  (red) to 7.22 nm (purple).

The fully ON thick filament can therefore be switched OFF by decreasing filament stress to zero, further supporting the mechano-sensing hypothesis. The extent of switching OFF, measured by the intensities of the ML1 and M2 reflections ( $I_{ML1}$  and  $I_{M2}$ ; Fig. 2a, b) or  $S_{M6}$  (Fig. 2c), increases with a half-time of  $\sim 20$  ms during the zero-stress period. This is much slower than the decrease in force (continuous line) or the fraction of myosin motors attached to actin, estimated from sarcomere stiffness, which has a half-time of  $\sim 3$  ms (Extended Data Fig. 2)<sup>17</sup>. Thus the rate of switching OFF the thick filament at zero force is limited by a transition in the detached myosin motors. Biochemical kinetic modelling suggests that the rate-limiting step is a stress-sensitive isomerization between states with bound ATP hydrolysis products (see Supplementary Discussion and Extended Data Fig. 2).

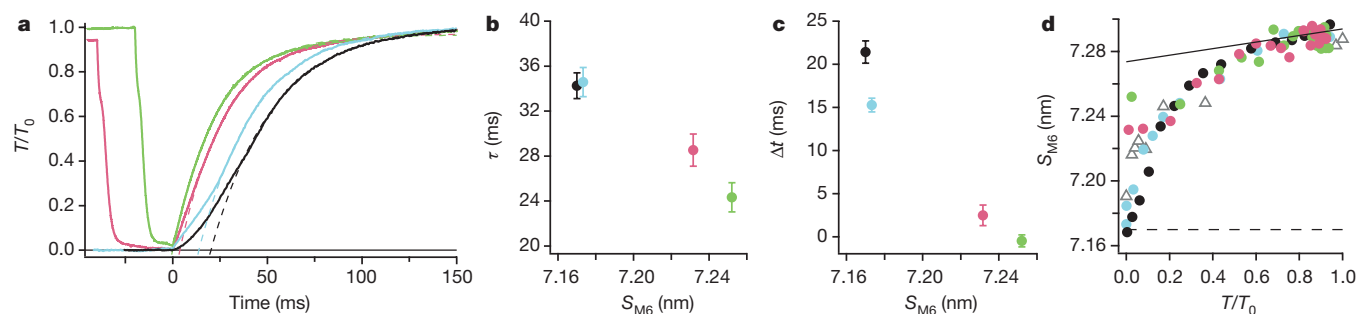
When force was allowed to redevelop after 40 ms at zero thick-filament stress, the filament returned to a fully ON structure, as measured by  $I_{ML1}$ ,  $I_{M2}$  and  $S_{M6}$ , again with a half-time of  $\sim 20$  ms, similar to that of force redevelopment (Fig. 2). Recovery of the intensity and interference fine structure of the M3 reflection had an  $\sim 10$ -ms lag, corresponding to force generation by disordered motors formed at zero stress (see Supplementary Discussion and Extended Data Fig. 3). The temporal correspondence between force redevelopment and recovery of the ON structure of the thick filament suggests that the rate of force generation in muscle is controlled by the availability of ON motors, and therefore by thick-filament structure.

To test the generality of this conclusion, we compared isometric force development after different periods of zero-force shortening to that at the start of calcium activation, in which force development starts from the end of the  $\sim 10$ -ms latent period (Fig. 3a, solid black line). The later time course of force development is well fit by an exponential (dashed black line) with a time constant ( $\tau$ ) of 34 ms (Fig. 3b, black circle), and a delay ( $\Delta t$ ) of 21 ms after the latent period (Fig. 3c). Force development after 20 ms of zero-force

shortening at the start of electrical stimulation (light blue) has the same  $\tau$  and a slightly smaller  $\Delta t$  (15 ms, measured from the end of shortening). In both cases, force develops from an almost fully OFF state of the thick filament, with  $S_{M6}$  close to its resting value (Fig. 3d). In contrast, force development after 40 ms of zero-force shortening from  $T_0$ , starting with the thick filament about half-OFF as



**Figure 2 | The OFF structure of the thick filament is transiently restored when the load on a fully active muscle fibre is removed.** Top, sarcomere length change, shown in nanometres per half-sarcomere (hs). **a–c**, Changes in  $I_{ML1}$  (**a**),  $I_{M2}$  (**b**) and  $S_{M6}$  (**c**) (circles) and force (solid line). Horizontal dashed lines, resting value of X-ray parameter; vertical dashed lines, time of end of shortening. X-ray data in **a** and **b** added from three fibres, those in **c** from two fibres. Filled/open circles denote data before/after the end of shortening.



**Figure 3 | The time course of force development is controlled by the regulatory state of the thick filament.** **a**, Time course of isometric force development following stimulation (black solid line); after 5% shortening from the latent period (light blue); and after 5% (green) or 10% (pink) shortening from the tetanus plateau, superimposed at the start of the force rise. **b**, **c**, Time constant ( $\tau$ ) (**b**) and delay ( $\Delta t$ ) (**c**) from exponential fits in **a** (mean  $\pm$  s.e.m.;  $n = 5$  fibres, except pink data ( $n = 3$ )) plotted against

$S_{M6}$  just before the start of force development. **d**, Instantaneous  $S_{M6}$ -force relationship for the four protocols in **a** (circles; data added from three fibres except pink and green ( $n = 2$ )). Continuous line, instantaneous filament compliance<sup>23</sup>; dashed line, resting value. Triangles, steady-state activation of demembrated mammalian muscle fibres at different free calcium concentrations (data from four fibre bundles). Colours in **b-d** denote protocols in **a**.

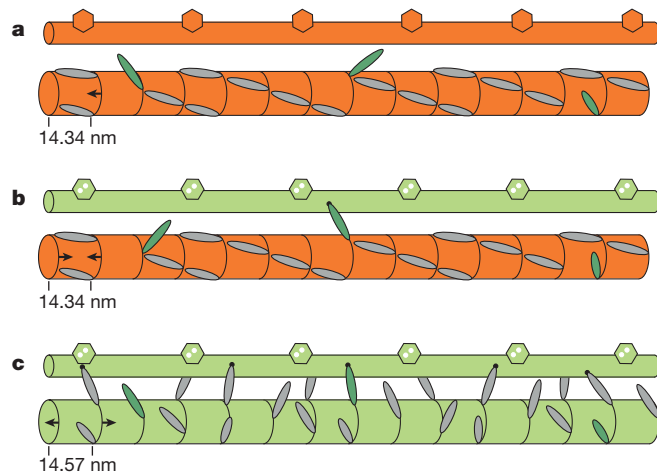
measured by  $S_{M6}$  (pink), is much faster ( $\tau = 28$  ms;  $\Delta t = 2$  ms). After only 20 ms of zero-force shortening from  $T_0$  (green), force development starts from a higher  $S_{M6}$  (Fig. 2c) and is even faster ( $\tau = 24$  ms;  $\Delta t = 0$  ms).

This correlation between the regulatory state of the thick filament and the kinetics of force development strongly supports the conclusion that the time course of force development is determined by availability of ON motors. For forces larger than  $0.1 T_0$ , there is a unique nonlinear relationship between  $S_{M6}$  and force in all four protocols (Fig. 3d). As the high-force part of this relationship is due to filament compliance (Fig. 3d, solid line), the thick filament may be considered to be fully ON for forces greater than about  $0.5 T_0$ . This relationship between thick-filament periodicity and stress is not limited to the protocols and muscle preparation used here; essentially the same relationship was observed when filament stress was modulated by varying the free calcium concentration in demembrated fibres from mammalian muscle in steady-state fixed-length conditions (Fig. 3d, triangles).

The results presented above lead to a novel dual-filament concept of muscle regulation (Fig. 4). If the external load is zero when the thin filaments are activated by calcium (Fig. 4b), muscle shortens at its maximum velocity, which is the same at the end of the latent period and during full activation<sup>16,18</sup> (Fig. 1). Less than 5% of the myosin motors are required to drive unloaded shortening<sup>17</sup>, and we propose that these motors are constitutively ON (Fig. 4, dark green ellipses). Unloaded shortening is therefore solely controlled by the regulatory state of the thin filament, and can be mobilized immediately after calcium release at low metabolic cost. The constitutively ON motors that drive unloaded shortening may be in the less ordered regions of the thick filament where MyBP-C is absent<sup>6,19</sup>, or a subset of the motors in the MyBP-C-containing region (the C zone). In either case, in the presence of an external load, the constitutively ON motors generate stress in the thick filament (Fig. 4c), releasing the remaining motors from the OFF state, and allowing development of the full isometric force  $T_0$ .

The molecular interactions stabilizing the OFF state of the myosin motors on the surface of the thick filaments in vertebrate skeletal muscle are not well understood, but may involve the motors binding to the myosin tail<sup>20</sup>, MyBP-C<sup>21</sup> and titin<sup>22</sup>, which all share the  $\sim 43$ -nm periodicity in the C zone. We postulate that these interactions stabilize the  $14.34 = 2 \times 7.17$ -nm axial periodicity of the helical OFF state (Fig. 4a, b), and that in their absence the thick filament reverts to the longer  $14.57$ -nm periodicity (Fig. 4c), determined by the packing of myosin tails in its backbone. Filament stress controls the transition between these two states by breaking the molecular interactions that stabilize the short or OFF state and decreasing the relative free energy of the long or ON state.

This thick-filament mechanosensing mechanism for the regulation of force generation in skeletal muscle has several wider implications. The OFF state of the thick filament inhibits ATP utilization by the great majority of the myosin motors, inducing the 'super-relaxed' state<sup>7</sup>, but the small fraction of constitutively ON motors allows the muscle to respond immediately to calcium activation when the external load is low. At high load, stress-dependent activation of the thick filament initiates a positive feedback loop, mobilizing more motors that generate more force to further increase thick-filament stress. This phase of motor unpacking dominates the initial accelerating phase of force development after calcium activation (Fig. 3a), consistent with the dependence of the kinetics of force development on thick-filament structure (Fig. 3b, c). Thus, the physiological rate of force development is determined by the time course of thick-filament activation. The classical force-velocity relationship of skeletal muscle, the fundamental determinant of its dynamic performance and efficiency, also has a new molecular explanation in terms of thick-filament mechanosensing. Finally, as thick-filament structure and protein composition are essentially the same in heart and skeletal muscle, thick-filament mechanosensing may also be a



**Figure 4 | Dual-filament regulation in skeletal muscle.** **a**, Resting muscle. **b**, **c**, Calcium activation at low and high load, respectively. Thin (top structure) and thick (bottom structure) filaments may be OFF (orange) or ON (green). The OFF thick filament has a compressed  $14.34$ -nm backbone periodicity (arrows); most myosin motors (grey) form helical tracks, but others (dark green) are constitutively ON. When the thin filament is activated by calcium binding to troponin (hexagons), the constitutively ON motors drive filament sliding at low load (**b**). At high load (**c**), force generated by constitutively ON motors switches the thick filament ON ( $14.57$ -nm periodicity), releasing the remaining motors.

fundamental component of the regulation of contractility in the heart, opening the possibility of new approaches for therapeutic control of cardiac output.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 23 April; accepted 21 September 2015.**

**Published online 11 November 2015.**

- Gordon, A. M., Homsher, E. & Regnier, M. Regulation of contraction in striated muscle. *Physiol. Rev.* **80**, 853–924 (2000).
- Huxley, H. E. & Brown, W. The low-angle X-ray diagram of vertebrate striated muscle and its behaviour during contraction and rigor. *J. Mol. Biol.* **30**, 383–434 (1967).
- Reconditi, M. *et al.* Motion of myosin head domains during activation and force development in skeletal muscle. *Proc. Natl Acad. Sci. USA* **108**, 7236–7240 (2011).
- Haselgrove, J. C. X-ray evidence for conformational changes in the myosin filaments of vertebrate striated muscle. *J. Mol. Biol.* **92**, 113–143 (1975).
- Woodhead, J. L. *et al.* Atomic model of a myosin filament in the relaxed state. *Nature* **436**, 1195–1199 (2005).
- Luther, P. K. *et al.* Direct visualization of myosin-binding protein C bridging myosin and actin filaments in intact muscle. *Proc. Natl Acad. Sci. USA* **108**, 11423–11428 (2011).
- Stewart, M. A., Franks-Skiba, K., Chen, S. & Cooke, R. Myosin ATP turnover rate is a mechanism involved in thermogenesis in resting skeletal muscle fibers. *Proc. Natl Acad. Sci. USA* **107**, 430–435 (2010).
- Linari, M. *et al.* Interference fine structure and sarcomere length dependence of the axial x-ray pattern from active single muscle fibers. *Proc. Natl Acad. Sci. USA* **97**, 7226–7231 (2000).
- Reconditi, M. *et al.* Sarcomere-length dependence of myosin filament structure in skeletal muscle fibres of the frog. *J. Physiol. (Lond.)* **592**, 1119–1137 (2014).
- Irving, T. *et al.* Thick-filament strain and interfilament spacing in passive muscle: effect of titin-based passive tension. *Biophys. J.* **100**, 1499–1508 (2011).
- Caputo, C., Edman, K. A., Lou, F. & Sun, Y. B. Variation in myoplasmic  $\text{Ca}^{2+}$  concentration during contraction and relaxation studied by the indicator fluo-3 in frog muscle fibres. *J. Physiol. (Lond.)* **478**, 137–148 (1994).
- Yagi, N. An X-ray diffraction study on early structural changes in skeletal muscle contraction. *Biophys. J.* **84**, 1093–1102 (2003).
- Rome, E., Offer, G. & Pepe, F. A. X-ray diffraction of muscle labelled with antibody to C-protein. *Nat. New Biol.* **244**, 152–154 (1973).
- Reconditi, M. *et al.* The myosin motor in muscle generates a smaller and slower working stroke at higher load. *Nature* **428**, 578–581 (2004).
- Huxley, H., Reconditi, M., Stewart, A. & Irving, T. X-ray interference studies of crossbridge action in muscle contraction: evidence from quick releases. *J. Mol. Biol.* **363**, 743–761 (2006).
- Brunello, E. *et al.* Structural changes in the myosin filament and cross-bridges during active force development in single intact frog muscle fibres: stiffness and X-ray diffraction measurements. *J. Physiol. (Lond.)* **577**, 971–984 (2006).
- Piazzesi, G. *et al.* Skeletal muscle performance determined by modulation of number of myosin motors rather than motor force or stroke size. *Cell* **131**, 784–795 (2007).
- Lombardi, V. & Menchetti, G. The maximum velocity of shortening during the early phases of the contraction in frog single muscle fibres. *J. Muscle Res. Cell Motil.* **5**, 503–513 (1984).
- Bennett, P. M. & Gautel, M. Titin domain patterns correlate with the axial disposition of myosin at the end of the thick filament. *J. Mol. Biol.* **259**, 896–903 (1996).
- Jung, H. S., Komatsu, S., Ikebe, M. & Craig, R. Head-head and head-tail interaction: a general mechanism for switching off myosin II activity in cells. *Mol. Biol. Cell* **19**, 3234–3242 (2008).
- Pfuhl, M. & Gautel, M. Structure, interactions and function of the N-terminus of cardiac myosin binding protein C (MyBP-C): who does what, with what, and to whom? *J. Muscle Res. Cell Motil.* **33**, 83–94 (2012).
- Muhle-Goll, C. *et al.* Structural and functional studies of titin's fn3 modules reveal conserved surface patterns and binding to myosin S1—a possible role in the Frank-Starling mechanism of the heart. *J. Mol. Biol.* **313**, 431–447 (2001).
- Brunello, E. *et al.* The contributions of filaments and cross-bridges to sarcomere compliance in skeletal muscle. *J. Physiol. (Lond.)* **592**, 3881–3899 (2014).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank M. Dolfi and J. Gorini for electronic and mechanical engineering support and P. Panine for assistance at the beamline. We thank ESRF for beamtime, and Ente Cassa di Risparmio di Firenze 2010.1402, FIRB-Futuro in Ricerca project RBFR08JAMZ, MIUR-PRIN project 2010R8JK2X (Italy), MRC (UK) and ESRF for financial support.

**Author Contributions** M.L., E.B., M.R., L.F., M.C., G.P., V.L. and M.I. contributed to the conception and design of the experiments, the collection, analysis and interpretation of data, and drafting or critical revision of the article. T.N. contributed to data collection and analysis.

**Additional Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to V.L. ([vincenzo.lombardi@unifi.it](mailto:vincenzo.lombardi@unifi.it)).



## METHODS

No statistical methods were used to predetermine sample size.

**Muscle fibres and experimental protocol.** Adult male frogs (*Rana temporaria*, 3–6 years old) were killed by decapitation and destruction of the brain and spinal cord, as approved by the Ethical Committee for the animal experimentation of the University of Florence, following European Community Council Directive 86/609/EEC. Frogs were chosen at random for each experiment from the available batch, and each experiment was performed using protocols with internal controls on the same frog muscle fibre to minimize systematic errors owing to inter-animal variation. Single fibres ~6 mm long were dissected from the tibialis anterior muscle and mounted via aluminium foil clips attached to the tendons in a trough containing Ringer's solution (115 mM NaCl, 2.5 mM KCl, 1.8 mM CaCl<sub>2</sub> and 3 mM phosphate buffer, pH 7.1). A pair of mica windows was positioned close to the fibre, about 600 µm apart, to minimize the X-ray path in solution. Sarcomere length, fibre length and cross-sectional area were measured with a 40× water immersion objective and a 25× eyepiece. Resting sarcomere length was set to  $2.14 \pm 0.02$  µm (mean  $\pm$  s.d.). Force was measured with a capacitance transducer<sup>24</sup>. The length of a population of sarcomeres in a 1–2 mm segment of the fibre was measured with a striation follower<sup>25</sup> in control tetani before mounting the fibre vertically at the beamline. Trains of stimuli of alternating polarity, at a frequency of 20–24 Hz, were delivered every 4 min at 4 °C via platinum electrodes on the top and bottom edges of the opposing windows to elicit fused tetani lasting 390–550 ms. A steady shortening of  $5.0 \pm 0.2\%$  (mean  $\pm$  s.e.m.,  $n = 8$ ) of the initial fibre length ( $L_0$ ) at the maximum velocity of shortening ( $V_0$ ) was applied either 5 ms after the first stimulus, to keep the force at zero, or at the plateau of the isometric tetanus ( $T_0$ ). After the imposed shortening, the isometric force developed at the new sarcomere length (Figs 1a and 3a). Alternatively, the fibre was slowly stretched at rest by 5%  $L_0$  and a steady shortening of 10%  $L_0$  ( $10.0 \pm 0.3\%$ ,  $n = 5$ ) at  $V_0$  was imposed at the plateau of an isometric tetanus. This protocol minimized sarcomere length differences at the end of ramp shortening between the various protocols used for X-ray measurements.  $V_0$  measured with the striation follower was  $2.7 \pm 0.2$  µm s<sup>-1</sup> per half-sarcomere (mean  $\pm$  s.d.). Stiffness during 10% shortening at  $V_0$  (Extended Data Fig. 2) was measured in fibres isolated from tibialis anterior muscles of adult male *Rana esculenta* at the Laboratory of Physiology (University of Florence, Italy), using the striation follower to monitor the half-sarcomere length change in response to step stretches in a segment near the force transducer end of the fibre.

Permeabilized muscle fibre segments were prepared from the psoas muscle of adult (~18 weeks old) male New Zealand white rabbits as described previously<sup>26</sup>. Bundles of 5–6 fibres, about 3–4 fibres wide and 1–2 fibres deep, were dissected on the day of the experiment and mounted horizontally in relaxing solution at ~2.4 µm sarcomere length between the lever arms of a strain gauge force transducer and a loudspeaker motor on a mechanical apparatus<sup>26</sup> modified for synchrotron X-ray experiments. Before each experiment, the ends of the bundle were fixed with glutaraldehyde and glued to aluminium foil clips with shellac dissolved in ethanol. Relaxing solution contained: 100 mM TES buffer, 7.7 mM MgCl<sub>2</sub>, 25 mM EGTA, 5.4 mM Na<sub>2</sub>ATP, 19.1 mM Na<sub>2</sub>-creatine phosphate (CP) and 10 mM reduced glutathione (GSH). Pre-activating solution contained: 100 mM TES, 6.9 mM MgCl<sub>2</sub>, 0.1 mM EGTA, 24.9 mM HDTA, 5.5 mM Na<sub>2</sub>ATP, 19.5 mM Na<sub>2</sub>CP and 10 mM GSH. Activating solution contained: 100 mM TES, 6.8 mM MgCl<sub>2</sub>, 25 mM CaEGTA, 5.5 mM Na<sub>2</sub>ATP, 19.5 mM Na<sub>2</sub>CP and 10 mM GSH. All solutions had 5 mM Mg-ATP; 1.2 mM free Mg<sup>2+</sup>; 199 mM ionic strength; pH 7.1 at 25 °C. Relaxing and activating solutions were mixed to obtain a series of partial activating solutions with the required free calcium ion concentrations. The osmotic agent dextran T500 (5% w/v) was added to all experimental solutions to reduce the interfibrillar spacing to a value similar to that of intact muscle (ref. 26 and references therein). The bundle was activated in a multidrop apparatus using a temperature-jump technique<sup>26</sup>; it was kept in pre-activating solution at low temperature (1 °C) for 2 min, then transferred to activating solution at 1 °C, in which little force was developed. When this force became steady (within 10 s) the bundle was transferred to activating solution at 25 °C and, following full force development, it was transferred to air for the X-ray exposure.

The investigators were not blinded to allocation during experiments and outcome assessment.

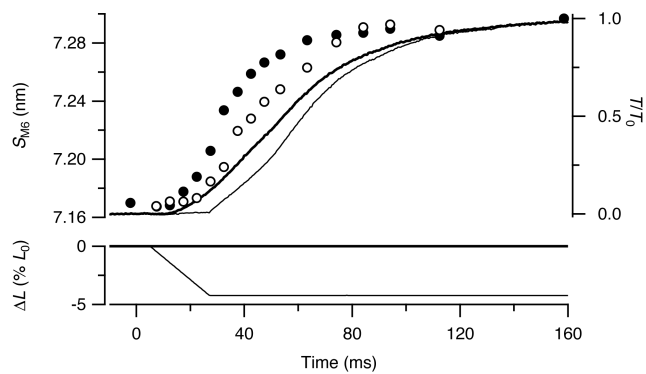
**X-ray data collection.** For intact fibres, the trough was sealed to prevent solution leakage, and the fibre was mounted vertically at beamline ID2 of the European Synchrotron Radiation Facility (ESRF)<sup>27</sup>, which provided up to  $2 \times 10^{13}$  photons per second at 0.1 nm wavelength in a beam of size ~300 µm (horizontal, full width at half-maximum (FWHM)) and ~100 µm (vertical) at the fibre. The beam was attenuated for fibre alignment. To minimize radiation damage, X-ray exposure was limited to the data collection period using a fast

electromagnetic shutter (nmLaser Products, Inc.) and the fibre was moved vertically by 100–200 µm between tetani. Data were collected from 50–80 tetani in each fibre with no detectable sign of radiation damage. X-ray diffraction patterns were recorded using the FReLoN charge-coupled device (CCD)-based detector with  $2,048 \times 2,048$  pixels binned by 16 in the horizontal direction and 2 in the vertical direction before the readout to increase the signal-to-noise ratio. Time frames of 3.5–6.5 ms were collected at rest, during 5%  $L_0$  steady shortening at  $V_0$  imposed 5 ms after the first stimulus, at the plateau of the isometric tetanus, during 5 or 10%  $L_0$  steady shortening at  $V_0$  imposed at the plateau of an isometric tetanus, and during force development following such shortening. The camera length was 10 m for recording the interference fine structure of the myosin-based M3 reflection and 3 m for recording the meridional reflections up to the M6. X-ray data are presented from ten fibres (five at 3 m and five at 10 m camera length, each fibre from a different frog) with a cross-sectional area of  $20,000 \pm 6,000$  µm<sup>2</sup> (mean  $\pm$  s.d.) and isometric plateau force ( $T_0$ ) of  $254 \pm 90$  kPa.

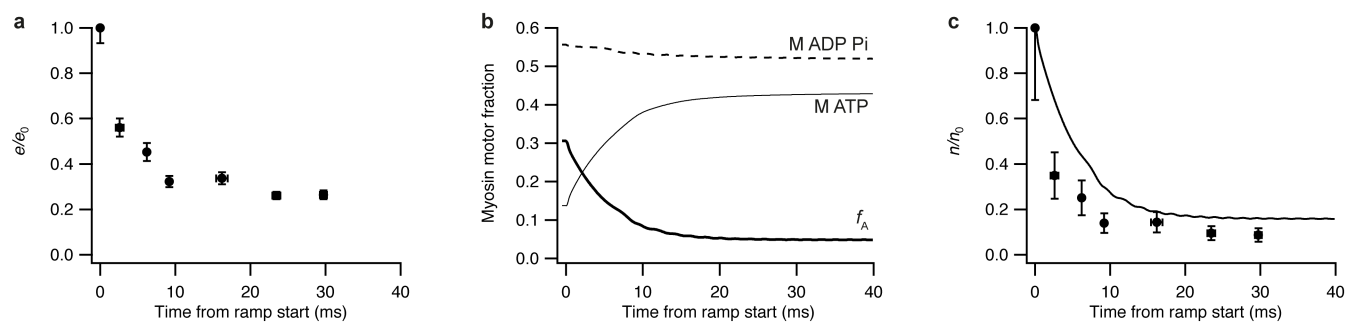
For X-ray experiments on permeabilized muscle fibres from rabbit psoas muscle, the multi-drop apparatus, which provided rapid solution exchange, temperature jumps and X-ray exposure in air, was mounted horizontally at the beamline. Data from 2–4 bundles were added for each point in Fig. 3d to increase the signal-to-noise ratio. Average bundle width was  $260 \pm 15$  µm; average fibre diameter was  $74 \pm 4$  µm (mean  $\pm$  s.d.).

**X-ray data analysis.** X-ray diffraction data were analysed using the SAXS package (P. Boescke, ESRF), Fit2D (A. Hammersley, ESRF) and IgorPro (WaveMetrics, Inc.). Two-dimensional patterns were centred and aligned using the equatorial 1,1 reflections, then mirrored horizontally and vertically. Data from fibres collected either at 10 m (2–5 fibres) or 3 m (2–5 fibres) were added together. The number of fibres was chosen in order to measure changes in relative intensities and spacings of the relevant X-ray reflections with adequate signal-to-noise; this could be achieved with a small number of fibres because the X-ray signals can be measured with extremely high precision and low biological variability<sup>3,8,9</sup>; for example, the spacing of the M6 reflection has an experimental standard deviation, including measurement and biological variability, of 0.005 nm (ref. 8), which is about 4% of the difference under investigation, between the resting state and  $T_0$ . The distribution of diffracted intensity along the meridional axis of the X-ray pattern (parallel to the fibre axis) was calculated by integrating from  $0.0046$  nm<sup>-1</sup> on either side of the meridian for the myosin-based M1 and M2 reflections,  $0.012$  nm<sup>-1</sup> for the M3 reflection and  $0.019$  nm<sup>-1</sup> for the M6 reflection. The first myosin layer line (ML1) was integrated in the region between  $0.064$  and  $0.037$  nm<sup>-1</sup> from the meridional axis. Background intensity distributions were fitted using a convex hull algorithm and subtracted; the small background remaining when the convex hull algorithm had been used was removed using the intensity from a nearby region of the X-ray pattern containing no reflections. Integrated intensities were obtained from the following axial regions: M1,  $0.021$ – $0.024$  nm<sup>-1</sup>; M2,  $0.046$ – $0.048$  nm<sup>-1</sup>; M3,  $0.067$ – $0.072$  nm<sup>-1</sup>; M6,  $0.133$ – $0.144$  nm<sup>-1</sup>; and ML1,  $0.019$ – $0.023$  nm<sup>-1</sup>. The limits for ML1 were chosen to exclude the contribution of the first actin layer line. The cross-meridional width of the M1, M2, M3 and M6 reflections was determined from the integrated intensity in a zone parallel to the equatorial axis in the axial regions specified above for the four reflections using a Gaussian fit across the meridian in the regions  $\pm 0.0037$  nm<sup>-1</sup>,  $\pm 0.0073$  nm<sup>-1</sup>,  $\pm 0.018$  nm<sup>-1</sup> and  $\pm 0.034$  nm<sup>-1</sup>, respectively. The interference components of the M3 reflection were determined by fitting multiple Gaussian peaks with the same axial width to the meridional intensity distribution, and the total intensity of the reflection was calculated as the sum of the component peaks. The spacing of each reflection was determined from the weighted mean of the component peaks, and calibrated using an M3 spacing of 14.34 nm in the resting fibre<sup>4</sup>. The combined instrumental point spread function was negligible compared with the radial width of the M3 reflection. Force, stimulus, sarcomere length, fibre length change and X-ray acquisition timing were collected and analysed using LabVIEW (National Instruments).

24. Huxley, A. F. & Lombardi, V. A sensitive force transducer with resonant frequency 50 kHz. *J. Physiol. (Lond.)* **305**, 15–6P (1980).
25. Huxley, A. F., Lombardi, V. & Peachey, L. D. A system for fast recording of longitudinal displacement of a striated muscle fibre. *J. Physiol. (Lond.)* **317**, 12P–13P (1981).
26. Linari, M., Caremani, M., Piperio, C., Brandt, P. & Lombardi, V. Stiffness and fraction of myosin motors responsible for active force in permeabilized muscle fibers from rabbit psoas. *Biophys. J.* **92**, 2476–2490 (2007).
27. Narayanan, T., Diat, O. & Boescke, P. SAXS and USAXS on the high brilliance beamline at the ESRF. *Nucl. Instrum. Methods Phys. Res. Sect. A* **467–468**, 1005–1009 (2001).



**Extended Data Figure 1 | The increase in the spacing of the M6 reflection on activation is delayed by imposing a period of unloaded shortening.** The top traces show  $S_{M6}$  (circles) superimposed on force (continuous line); filled/open circles and thicker/thinner line denote data from fixed-end tetani and tetani with imposed shortening, respectively. The bottom traces show imposed length change ( $\Delta L$ , expressed as percentage of initial fibre length (%  $L_0$ )). X-ray data added from one/two repeats of the protocol in three muscle fibres.

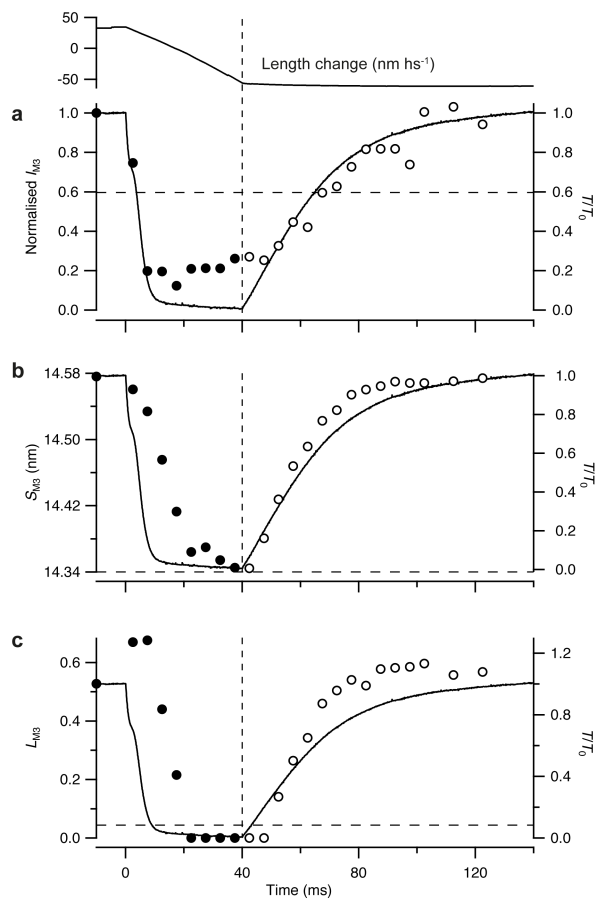


### Extended Data Figure 2 | Stiffness changes during unloaded shortening.

**a**, Half-sarcomere stiffness ( $e$ ) relative to that at the plateau of an isometric tetanus ( $e_0$ ), at different times after the start of unloaded shortening of 10%  $L_0$  applied at the tetanus plateau, calculated from the ratio of force and half-sarcomere length changes in response to 0.2%  $L_0$  step stretches complete in 100  $\mu$ s. Mean  $\pm$  s.e.m. from four fibres; species, *Rana esculenta*, 4  $^{\circ}$ C;  $e_0 = 0.27 \pm 0.01 T_0 \text{ nm}^{-1}$ ;  $T_0 = 137 \pm 13 \text{ kPa}$ . **b**, Fraction of myosin motors attached to actin ( $f_A$ ; thick solid line) and fractions with bound ATP (thin solid line) or ADP and inorganic phosphate ( $P_i$ ) (dashed

line) as a function of time during unloaded shortening, calculated from the kinetic model described in the Supplementary Discussion. **c**, Number of myosin motors attached to actin ( $n$ ) relative to that at the plateau of an isometric tetanus ( $n_0$ ), at different times during unloaded shortening, calculated from  $e/e_0$  in **a** as described in the Supplementary Discussion. Mean  $\pm$  s.e.m. from four fibres; s.e.m. includes the contribution of errors in measured values of filament and parallel elasticity. The thick line was calculated by normalizing  $f_A$  in **b** for its tetanus plateau value.





**Extended Data Figure 3 | Changes in the intensity, spacing and interference fine structure of the M3 reflection associated with a period of unloaded shortening.** Top, sarcomere length change in nanometres per half-sarcomere. **a–c**, Changes in intensity ( $I_{M3}$ ; **a**), spacing ( $S_{M3}$ ; **b**) and interference fine structure ( $L_{M3}$ ; **c**) (circles) superimposed on force (continuous line). Horizontal dashed lines, resting value of X-ray parameter; vertical dashed lines, time of end of shortening. X-ray data added from three fibres. Filled/open circles denote data before/after the end of shortening.

# Polarized endosome dynamics by spindle asymmetry during asymmetric cell division

Emmanuel Derivery<sup>1</sup>, Carole Seum<sup>1</sup>, Alicia Daeden<sup>1</sup>, Sylvain Loubéry<sup>1</sup>, Laurent Holtzer<sup>1</sup>, Frank Jülicher<sup>2</sup> & Marcos Gonzalez-Gaitan<sup>1</sup>

**During asymmetric division, fate determinants at the cell cortex segregate unequally into the two daughter cells. It has recently been shown that Sara (Smad anchor for receptor activation) signalling endosomes in the cytoplasm also segregate asymmetrically during asymmetric division<sup>1,2</sup>. Biased dispatch of Sara endosomes mediates asymmetric Notch/Delta signalling during the asymmetric division of sensory organ precursors in *Drosophila*<sup>1</sup>. In flies, this has been generalized to stem cells in the gut<sup>3</sup> and the central nervous system<sup>1</sup>, and, in zebrafish, to neural precursors of the spinal cord<sup>4</sup>. However, the mechanism of asymmetric endosome segregation is not understood. Here we show that the plus-end kinesin motor Klp98A targets Sara endosomes to the central spindle, where they move bidirectionally on an antiparallel array of microtubules. The microtubule depolymerizing kinesin Klp10A and its antagonist Patronin generate central spindle asymmetry. This asymmetric spindle, in turn, polarizes endosome motility, ultimately causing asymmetric endosome dispatch into one daughter cell. We demonstrate this mechanism by inverting the polarity of the central spindle by polar targeting of Patronin using nanobodies (single-domain antibodies). This spindle inversion targets the endosomes to the wrong cell. Our data uncover the molecular and physical mechanism by which organelles localized away from the cellular cortex can be dispatched asymmetrically during asymmetric division.**

We first identified Klp98A as the kinesin mediating Sara endosome motility during sensory organ precursor (SOP) division (Fig. 1). Klp98A is the *Drosophila* homologue of mammalian KIF16B, an early endosomal kinesin containing a phosphatidylinositol 3-phosphate-binding PX domain<sup>5</sup>. Indeed, Klp98A localizes to Sara-positive early endosomes (Extended Data Fig. 1a–e).

During SOP division, Klp98A–GFP-positive Sara endosomes segregate to the pIIa daughter, but not the pIIb<sup>1,2</sup> (Fig. 1a, Supplementary Video 1). Sara endosomes were monitored by following Delta 20 min after internalization (iDelta<sub>20</sub>) through an improved antibody internalization assay<sup>1</sup>. iDelta<sub>20</sub> parallels Sara endosome dynamics in the controls and mutants studied here (*in vivo* and primary cultures; Extended Data Fig. 2). Like KIF16B<sup>5,6</sup>, purified Klp98A (Fig. 1b) binds specifically to phosphatidylinositol 3-phosphate (Extended Data Fig. 1f) and is a plus-end-directed motor (Fig. 1c) whose velocity is  $0.76 \pm 0.02 \mu\text{m s}^{-1}$  (mean  $\pm$  s.e.m.,  $n = 345$  motility strides; Fig. 1d).

To study Klp98A function, we generated deletions within the motor domain (Klp98A<sup>Δ6</sup>, Klp98A<sup>Δ7</sup> and Klp98A<sup>Δ8</sup>, 6, 7 and 8-base-pair deletions, respectively) and a clean coding sequence deletion (Klp98A<sup>Δ47</sup>) (Extended Data Fig. 1a, b, g–k). Except Klp98A<sup>Δ6</sup>, all are protein nulls. In Klp98A<sup>−</sup>, Sara endosomes move diffusively (diffusion coefficient  $D = 0.0021 \pm 0.0001 \mu\text{m}^2 \text{s}^{-1}$ , mean  $\pm$  s.e.m.,  $n = 4$  independent methods each based on at least 100 tracks; Fig. 1e, g and Extended Data Fig. 4a–e; ‘Mean Square Displacement analysis’ in Methods). Therefore, Klp98A mediates Sara endosome motility.

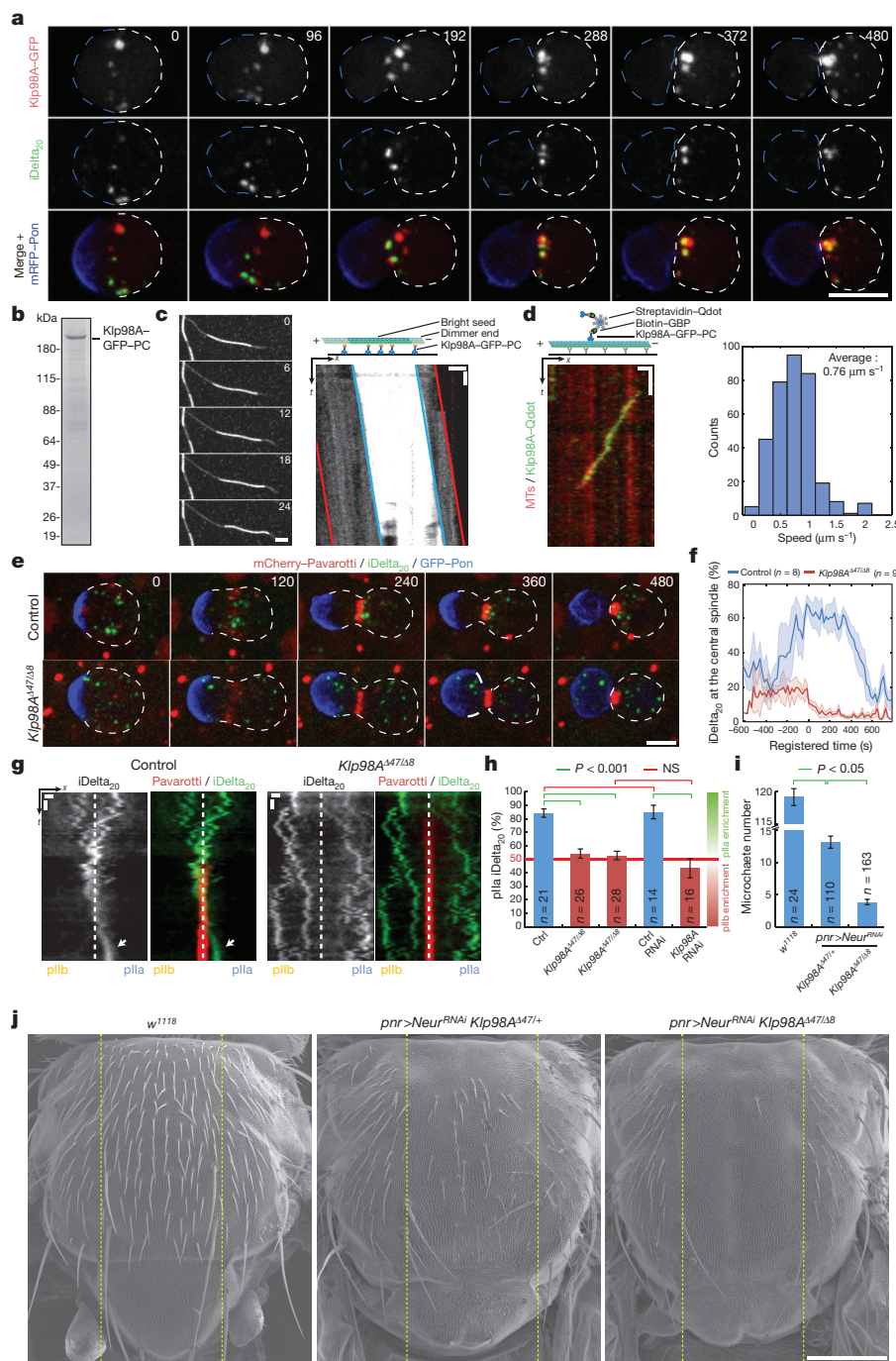
In wild-type cells, Sara endosomes move on microtubules to the Pavarotti-positive central spindle (Fig. 1e–g, Supplementary Video 2) and, late in cytokinesis, to pIIa (Fig. 1a, e, g (arrows) and h). Spindle microtubule plus-ends are oriented towards the equator<sup>7</sup>, explaining central spindle endosomal targeting by a plus-end motor. Indeed, Sara endosome central spindle targeting fails in Klp98A<sup>−</sup> mutants (Fig. 1e–g, Supplementary Video 2). Importantly, in Klp98A<sup>−</sup> mutants and upon RNAi-mediated Klp98A knockdown, endosomes are symmetrically dispatched (Fig. 1e, g, h).

Klp98A-mediated motility contributes to cell fate assignment through asymmetric Notch signalling, but this activity is redundantly covered by Neuralized and Numb<sup>8</sup>. Indeed, Klp98A<sup>−</sup>;pnr > neur<sup>RNAi</sup> double mutants show a synergistic fate assignment phenotype: the notum is largely void of bristles (Fig. 1i, j and Extended Data Fig. 3a, b; ‘Quantification of the Neur/Numb phenotypes’ in Methods). Conversely, Klp98A;Numb double mutants strongly suppress the diagnostic Numb<sup>−</sup> multiple socket phenotype<sup>9</sup> (Extended Data Fig. 3d–f). Therefore, having established the role of Klp98A motility in Notch signalling, we focus here on the mechanisms orchestrating asymmetric motility.

Central spindle targeting of Sara endosomes precedes asymmetric segregation to pIIa. We therefore focused on Sara endosome motility with respect to the central spindle reference frame. The central spindle is composed of the Pavarotti-positive core (containing antiparallel microtubules) plus the microtubules emanating from it<sup>10</sup> (Fig. 2a). We automatically tracked the Pavarotti core, defining a 2D cartesian reference frame whose origin is the Pavarotti centroid and whose  $x$  axis is the pIIb–pIIa axis (Fig. 2a, b and Supplementary Video 3; for algorithm and accuracy, Methods and Extended Data Fig. 4f–i). This also defines a Pavarotti width (PW) and length (PL, the length of the microtubule antiparallel array). We used the contracting PW for time-registration of our movie data sets (Fig. 2a–c; Methods and Extended Data Fig. 4o–x; registered time 0 represents anaphase B onset).

We then tracked Sara endosomes with respect to this reference frame (with 160 nm accuracy; Methods and Extended Data Fig. 4j–n). Automatic tracking and spatio-temporal registration provided a large data set (2,897 traces) from which a spatio-temporal density plot of endosomes at the central spindle was generated (Fig. 2d). For 500 s, endosomes remain mostly within the Pavarotti region (Fig. 2d). Remarkably, at the central spindle, motility along the  $x$  axis is bidirectional (Fig. 2e, Supplementary Video 4, Extended Data Fig. 4y). Motility along the  $y$  axis merely follows PW contraction (Fig. 2f), consistent with motility along central spindle microtubules, parallel to the  $x$  axis. Velocities are similar towards pIIa ( $0.18 \pm 0.1 \mu\text{m sec}^{-1}$ ; mean  $\pm$  s.e.m.,  $n = 422$  events) and pIIb ( $0.17 \pm 0.09 \mu\text{m sec}^{-1}$ ;  $n = 428$  events) and slower than *in vitro* (Fig. 1d), possibly due to crowding by microtubule-associated proteins<sup>11,12</sup>.

<sup>1</sup>Department of Biochemistry, Faculty of Sciences, University of Geneva, 30 Quai Ernest Ansermet, Geneva 1211, Switzerland. <sup>2</sup>Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Strasse 38, 01187 Dresden, Germany.



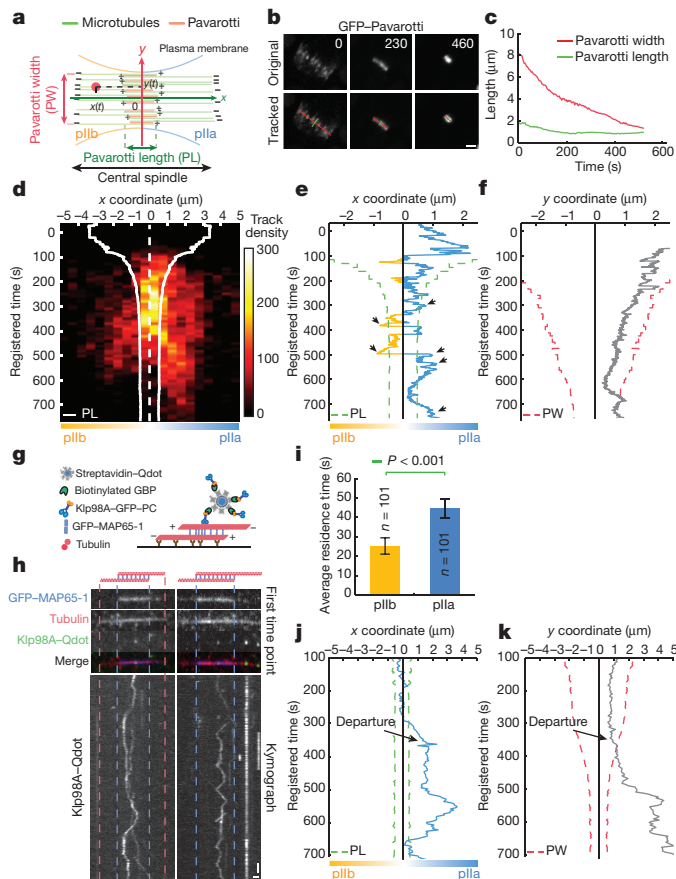
**Figure 1 | Klp98A controls Sara endosome motility and central spindle targeting.** **a**, Dividing SOP showing Klp98A-GFP, Sara endosomes (iDelta<sub>20</sub>) and mRFP-Pon (monomeric red fluorescent protein fused to the cortex localization domain of Pon, which labels the pIIB cortex: dashed blue lines). pIIB cortex, dashed white lines. 84 ± 3% iDelta<sub>20</sub> co-localizes with Klp98A (mean ± s.e.m.,  $n = 51$  cells with a total of 308 endosomes). **b**, Purified Klp98A (Coomassie SDS-PAGE; PC, Protein C tag). **c**, Left, gliding assay of polarity-marked microtubules on purified Klp98A. Right, kymograph. The short, dimmer minus-end leads; Klp98A is a plus-end motor. **d**, Left, kymograph of Klp98A-bound quantum dots (Qdots)

moving on microtubules (MTs). Right, speed distribution. **e–g**, iDelta<sub>20</sub> dynamics (**g**; kymograph of horizontal projection of **e**) and central spindle targeting in control and Klp98A<sup>Δ47/Δ8</sup> mutants (**f**). Time is registered between movies (see Methods). Registered time = 0, anaphaseB onset. **h**, iDelta<sub>20</sub> segregation after abscission. **i**, **j**, Scanning electron microscopy and microchaete numbers in *panier* region (dashed lines in **j**) in Klp98A/Neuralized mutants. In **h**, **i**, Kruskal–Wallis test. Scale bars: 5 μm (**a**, **e**), 1 μm/2 s (**c**), 5 μm/2 s (**d**), 1 μm/1 min (**g**), 200 μm (**j**).  $n$ , numbers of SOPs (**f**, **h**) or flies (**i**). Elapsed time, seconds. In **f**, **h**, **i** and throughout this report mean ± s.e.m. is shown.

Confinement within the Pavarotti region and bidirectional movement are both consistent with a plus-end motor switching direction on antiparallel microtubules. On single microtubules, Klp98A-bound quantum dots always maintain their directionality when resuming after a pause ( $n = 29$ ; Fig. 1d). We asked whether Klp98A could switch

direction in an antiparallel bundle. In an *in vitro* reconstitution assay, Klp98A-bound quantum dots move bidirectionally within antiparallel MAP65-1-mediated microtubule arrays<sup>13</sup> (Fig. 2g, h, Supplementary Video 5; 68% tracks ( $n = 150$ ) change direction after pausing). Therefore, Klp98A supports bidirectional motility in antiparallel arrays.





**Figure 2 | Asymmetric motility of Sara endosomes at the central spindle.** **a–c**, Central spindle reference frame (**a**), automatic tracking (**b**; PL, green; PW, red) and PW/PL dynamics (**c**). **d**, iDelta<sub>20</sub> spatio-temporal density plot at the central spindle. 2,897 tracks (45 cells) were registered in space/time and displayed in a single density plot. White line, averaged PL; dashed, Pavarotti centroid. **e, f**, Representative track on the *x* (**e**) and *y* (**f**) axes. **g, h**, *In vitro* reconstitution of bidirectional motility using Klp98A-bound quantum dots (Qdots) and antiparallel microtubule arrays organized by MAP65-1 (kymographs in **h**). **i**, pIla/pIib endosome residence time ( $n = 101$  tracks, Mann–Whitney test). **j, k**, *x* and *y* axis tracks; see departure event. Scale bars: 2  $\mu\text{m}$  (**b**); 1.2  $\mu\text{m}/2$  min (**h**).

Notably, *in vivo*, bidirectional endosome motility is asymmetric: the residence time in pIla is 1.8-fold longer than in pIib (Fig. 2i). Consistently, the spatio-temporal density plot is asymmetric (Fig. 2d). Furthermore, tracks overshoot beyond the Pavarotti region more frequently into pIla (Fig. 2e, arrows; see also Fig. 2d).

Eventually, endosomes depart from central spindle microtubules into the cytoplasm and therefore move also on the *y* axis (Fig. 2j, k). The longer pIla residence time and higher pIla overshoot frequency make this final departure asymmetric, explaining the biased segregation into pIla (Fig. 1h). Therefore, asymmetric endosome motility at the central spindle underlies asymmetric dispatch to pIla.

We then wondered whether the central spindle itself is asymmetric. Using Pavarotti spatio-temporal registration, we generated an ‘average cell’ to map the densities of the microtubule markers Jupiter<sup>14</sup> and SiR-tubulin<sup>15</sup> (microtubule markers), Patronin<sup>16</sup> (minus-end), and Pavarotti (plus-ends/antiparallel overlap) (Fig. 3a, b; Extended Data Fig. 5a, b; Supplementary Video 6). This ‘average cell’ reveals a polarity map of the central spindle consistent with electron microscopy reports<sup>17,18</sup>: plus-ends are in the middle and minus-ends on the outer side (Fig. 3c). Microtubule densities in general, and Patronin in particular, are ~20% higher on the pIib side (Fig. 3a, b; Extended Data Fig. 5c–i: quantifications and endogenous stainings). This asymmetry

depends on Par complex activity, and is absent in neighbouring cells dividing symmetrically (Extended Data Fig. 5d–k). We confirmed previous reports of asymmetric centrosomes in SOPs<sup>19–21</sup>, but this seems independent of central spindle or endosomal asymmetry (Extended Data Fig. 5l–s).

Microtubule asymmetry builds up during anaphaseB, concomitant with biased endosome motility, while, earlier, the metaphase spindle is symmetrical (Fig. 3d, e; Extended Data Figs 2i and 5l–n; Supplementary Video 7). During anaphaseB, the central spindle shrinks by microtubule depolymerization through depolymerizing kinesins like Klp10A<sup>22,23</sup>, among other factors<sup>24–26</sup>. Depolymerization dynamics are asymmetric: microtubule loss is faster in pIla (Fig. 3f). This could be explained by Patronin enrichment in the central spindle pIib outer side (Fig. 3a, b) where it binds to minus-ends, counteracting Klp10A-mediated depolymerization<sup>16,23,27,28</sup>.

Indeed, Klp10A/Patronin control asymmetric microtubule depolymerization: their depletion abolishes spindle asymmetry (Fig. 3e, g, h; Extended Data Fig. 6: controls, co-depletion and endogenous tubulin immunostainings). In *Patronin*-knockdown cells, both sides exhibit low microtubule densities characteristic of pIla (Fig. 3g, h), consistent with Patronin pIib enrichment in wild type (Fig. 3a, b) and its activity against depolymerization<sup>16,23,27,28</sup>. Conversely, upon knockdown of *Klp10A*, both sides exhibit high microtubule densities resembling pIib (Fig. 3g–i).

The parallelism between central spindle asymmetry and asymmetric endosome motility suggests that spindle asymmetry causes biased motility. Indeed, endosome motility at the central spindle and, therefore, segregation become symmetric in *Klp10A*- and *Patronin*-knockdowns, while early central spindle targeting is normal (Fig. 4a, b; Supplementary Video 8; Extended Data Fig. 7: controls and co-depletion). This uncovers a quantitative correlation between spindle and endosomal asymmetry (Fig. 4c).

Together, a plus-end motor (Fig. 1) and microtubule plus-ends facing the centre (Fig. 3a–c) explain why a higher pIib microtubule density (~20% enrichment) targets endosomes to pIla (~80% pIla, that is, >300% enrichment). In other words, endosomes move away from higher microtubule densities in pIib.

Based on a theoretical model (see Supplementary Equations) of plus-end endosomal motility on an antiparallel, asymmetric microtubule overlap (Fig. 4d), the steady-state endosome distribution is

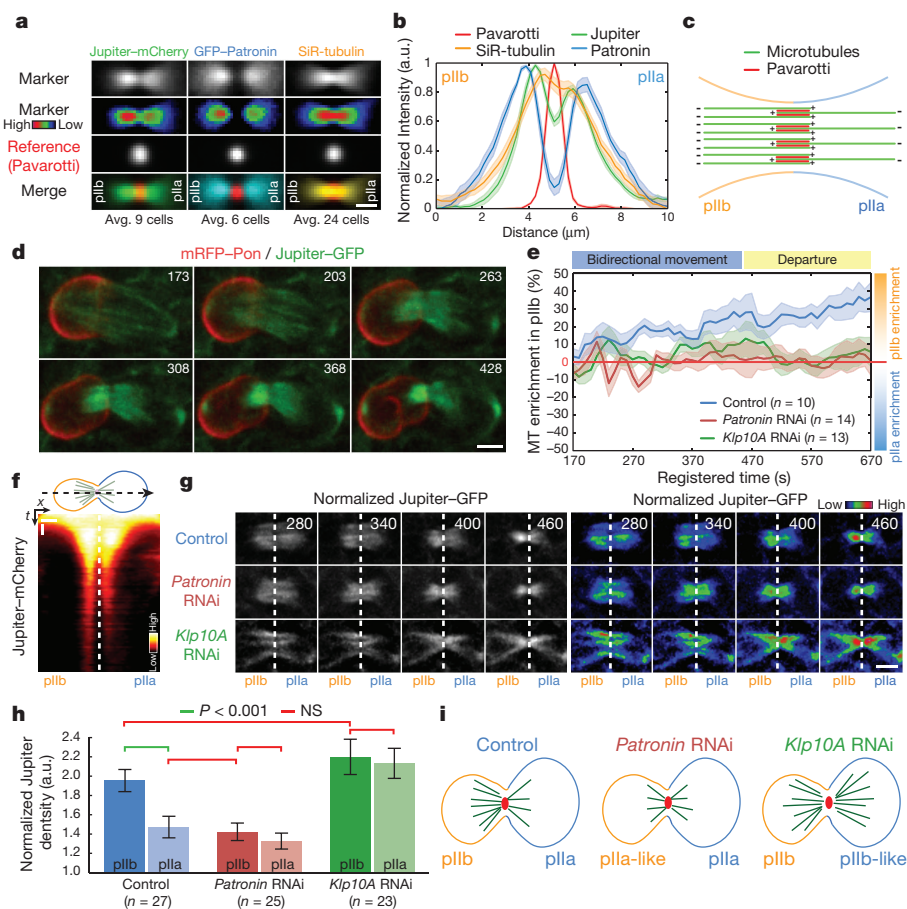
$$\frac{P_{\text{pIla}}}{P_{\text{pIib}}} = \frac{\rho_b}{\rho_a} e^{\frac{l k_{\text{on}} v (\rho_b - \rho_a)}{D k_{\text{off}}}} \quad (1)$$

Where  $P_{\text{pIla}}$ ,  $P_{\text{pIib}}$ , the probabilities for an endosome to be on either side of the antiparallel overlap;  $\rho_a$ ,  $\rho_b$ , microtubule densities in pIla/pIib, respectively;  $k_{\text{on}}$ ,  $k_{\text{off}}$ , microtubule association/dissociation constants of the motor, respectively;  $v$ , the endosome motor-driven velocity;  $D$ , the diffusion coefficient of endosomes detached from microtubules; and  $l$ , the antiparallel overlap length (Extended Data Fig. 10a; Supplementary Equations).

Based on equation (1), Fig. 4e shows how the pIla fraction of endosomes  $\frac{P_{\text{pIla}}}{P_{\text{pIla}} + P_{\text{pIib}}}$  depends on the normalized difference of microtubule

densities  $\Delta = \frac{\rho_b - \rho_a}{\rho_b + \rho_a}$ . We measured  $D = 0.0021 \pm 0.0001 \mu\text{m}^2 \text{s}^{-1}$

(Extended Data Fig. 4a–e),  $v = 0.173 \pm 0.007 \mu\text{m s}^{-1}$  (Fig. 2e),  $l = \text{PL} = 1 \pm 0.1 \mu\text{m}$  (Fig. 2d),  $k_{\text{off}} = 0.90 \pm 0.06 \text{s}^{-1}$  (Extended Data Fig. 10c) and  $k_{\text{on}} \rho = 0.05 \pm 0.01 \text{s}^{-1}$  ( $\rho$  corresponds to the average microtubule density, see Supplementary Information equation (36); Extended Data Fig. 10d). With these parameters, according to equation (1), a 20% microtubule pIib enrichment is amplified into 300% endosome pIla enrichment (Fig. 4e). Furthermore, if spindle asymmetry is inverted, endosomes become enriched in pIib.



**Figure 3 | The central spindle of SOPs is asymmetric.** **a**, Average densities of microtubule markers at the central spindle (live; late cytokinesis, registered time  $\sim 600$  s). **b**, Line scan through the average spindles in **a**. a.u., arbitrary units. **c**, Central spindle microtubule topology. **d**, Microtubule dynamics in a dividing SOP. **e**, Dynamics of central spindle microtubule (MT) pIIb enrichment ( $100 \times \frac{(\rho_b - \rho_a)}{\rho_a}$ ;  $\rho_b$ , pIIb microtubule density;  $\rho_a$ , pIIa) in indicated

genotypes. **f**, Microtubule density kymograph at central spindle. **g**, **h**, Jupiter-GFP in dividing SOPs of indicated genotypes and quantification (ANOVA). Normalization to pIIa centrosome. NS, not significant. **i**, Patronin/Klp10A phenotypes. Scale bars:  $2 \mu\text{m}$  (**a**, **e**, **g**),  $2 \mu\text{m}/1 \text{ min}$  (**f**). Elapsed time in **d** and **g** is given in seconds and corresponds to registered time;  $n$ , SOP number.

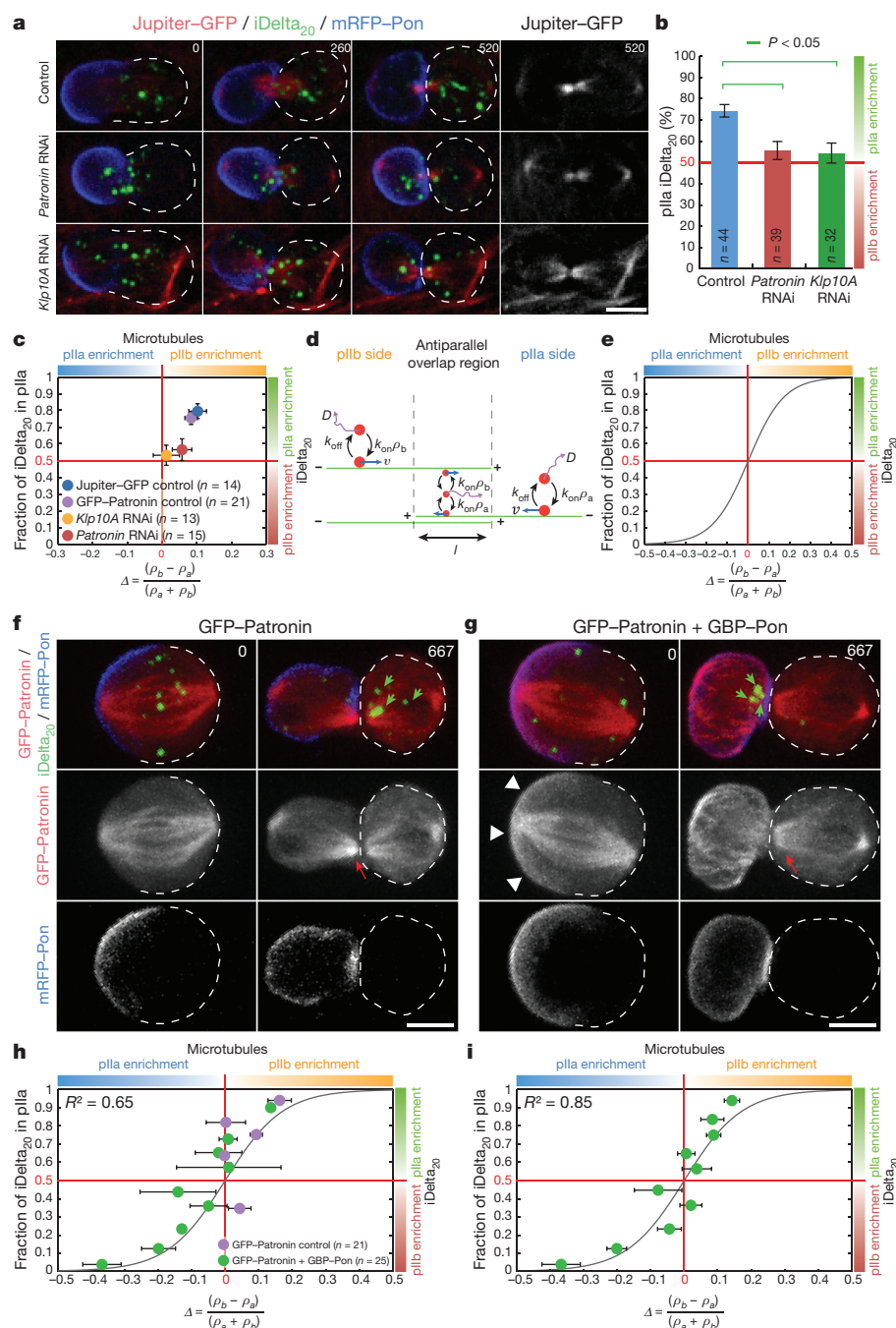
To generate this inverted spindle, we established a ‘nanobody assay’ based on GFP-binding-peptide (GBP)-Pon, a nanobody fused to the Pon localization domain (Supplementary Information; Extended Data Figs 8 and 9). GBP-Pon traps GFP-Patronin away from the spindle at the pIIb cortex thereby reducing, specifically in pIIb, Patronin-dependent protection against central spindle depolymerization (Fig. 4f, g, white arrowheads; Supplementary Video 9). This inverts spindle asymmetry (Fig. 4g, red arrows), which consequently inverts endosomal asymmetry (green arrows; Extended Data Fig. 8c, d, h for controls). SiR-Tubulin and endogenous acetyl-tubulin stainings confirmed this spindle inversion (Extended Data Figs 8e–g and 9h; Supplementary Video 10).

Interestingly, this assay generates a phenotypic series of different levels of spindle reversal and their corresponding endosomal reversals. These data fall on the theoretical curve obtained with independently measured parameters: equation (1) captures the observed spindle/endosome correlation (Fig. 4h; Extended Data Fig. 10f; Supplementary Information equation (36)). Beyond the nanobody assay, equation (1) accounts for our entire data set (Jupiter-GFP and GFP-Patronin controls, *Klp10A* and *Patronin* RNAi knockdowns, and nanobody assay; Fig. 4i; Extended Data Fig. 10g). Therefore our results

uncover the quantitative dependence of asymmetric endosome targeting on spindle asymmetry.

Here we identified *Klp10A*/Patronin as the machinery generating spindle asymmetry, which is read out by *Klp98A* to achieve asymmetric targeting of signalling endosomes. Asymmetric endosomal targeting contributes in turn to asymmetric cell fate assignment, confirming previous reports in flies<sup>1–3</sup> and fish<sup>4</sup>. Our data thus uncover a mechanism by which intracellular cargoes in general, and signalling endosomes in particular, can be targeted to one of the daughter cells during asymmetric division.

How could then other cargoes segregate symmetrically, if the spindle is asymmetric? Asymmetric targeting would only be efficient if  $k_{\text{on}}$ ,  $k_{\text{off}}$  and  $v$  are optimized to amplify the mild asymmetry of the spindle, otherwise concealed by noise sources in the cell. More generally, plus- and minus-end motors are present simultaneously in the same vesicle and thereby may counteract each other to achieve symmetrical dispatch (a sort of ‘tug of war’). Therefore, the precise landscape of microtubule polarity trails combined with the right cocktail of motors in vesicles provides the plasticity required to generate the plethora of molecular spatial patterns observed in polarized cells.



**Figure 4 | Central spindle asymmetry orchestrates Sara endosome asymmetry.** **a, b**, Microtubule/iDelta<sub>20</sub> dynamics (**a**) and iDelta<sub>20</sub> percentage in pIIa after abscission in SOPs of indicated genotypes (**b**; ANOVA). **c**, iDelta<sub>20</sub> pIIa fraction versus normalized microtubule pIIb enrichment ( $\Delta = \frac{\rho_b - \rho_a}{\rho_a + \rho_b}$ ) measured simultaneously in indicated genotypes (registered time  $\approx 600$  s). **d, e**, Endosome segregation model by motor transport on an antiparallel microtubule overlap (**d**) and endosome-segregation/spindle-asymmetry dependence from the model

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 19 December 2014; accepted 11 November 2015.

- Coumilleau, F., Furthauer, M., Knoblich, J. A. & Gonzalez-Gaitan, M. Directional Delta and Notch trafficking in Sara endosomes during asymmetric cell division. *Nature* **458**, 1051–1055 (2009).
- Loubéry, S. *et al.* Uninflatable and Notch control the targeting of Sara endosomes during asymmetric division. *Curr. Biol.* **24**, 2142–2148 (2014).

(**e**; Supplementary Information equation (36)). **f, g**, GFP-Patronin and iDelta<sub>20</sub> dynamics without (**f**) and with (**g**) expression of GBP-Pon. **h, i**, iDelta<sub>20</sub> pIIa fraction versus normalized microtubule pIIb enrichment (as in **c**) in indicated genotypes and theoretical curve (as in **e**). Data point binning performed according to iDelta<sub>20</sub> segregation (un-binned, Extended Data Fig. 10f). **i**, Plot, as in **h**, for the pooled data from all genotypes in this study. (Un-binned, Extended Data Fig. 10g). Scale bars, 5  $\mu$ m; elapsed time is given in seconds; *n*, SOP number.

- Montagne, C. & Gonzalez-Gaitan, M. Sara endosomes and the asymmetric division of intestinal stem cells. *Development* **141**, 2014–2023 (2014).
- Kressmann, S., Campos, C., Castanon, I., Furthauer, M. & Gonzalez-Gaitan, M. Directional Notch trafficking in Sara endosomes during asymmetric cell division in the spinal cord. *Nature Cell Biol.* **17**, 333–339 (2015).
- Hoepfner, S. *et al.* Modulation of receptor recycling and degradation by the endosomal kinesin KIF16B. *Cell* **121**, 437–450 (2005).
- Soppina, V. *et al.* Dimerization of mammalian kinesin-3 motors results in superprocessive motion. *Proc. Natl Acad. Sci. USA* **111**, 5562–5567 (2014).
- Glötzter, M. The 3Ms of central spindle assembly: microtubules, motors and MAPs. *Nature Rev. Mol. Cell Biol.* **10**, 9–20 (2009).



8. Fürthauer, M. & Gonzalez-Gaitan, M. Endocytic regulation of notch signalling during development. *Traffic* **10**, 792–802 (2009).
9. Wang, S., Younger-Shepherd, S., Jan, L. Y. & Jan, Y. N. Only a subset of the binary cell fate decisions mediated by Numb/Notch signaling in *Drosophila* sensory organ lineage requires Suppressor of Hairless. *Development* **124**, 4435–4446 (1997).
10. Fededa, J. P. & Gerlich, D. W. Molecular control of animal cell cytokinesis. *Nature Cell Biol.* **14**, 440–447 (2012).
11. Korten, T. & Diez, S. Setting up roadblocks for kinesin-1: mechanism for the selective speed control of cargo carrying microtubules. *Lab Chip* **8**, 1441–1447 (2008).
12. LaPointe, N. E. *et al.* The amino terminus of tau inhibits kinesin-dependent axonal transport: implications for filament toxicity. *J. Neurosci. Res.* **87**, 440–451 (2009).
13. Stoppin-Mellet, V., Fache, V., Portran, D., Martiel, J. L. & Vantard, M. MAP65 coordinate microtubule growth during bundle formation. *PLoS ONE* **8**, e56808 (2013).
14. Karpova, N., Bobinnec, Y., Fouix, S., Huitorel, P. & Debec, A. Jupiter, a new *Drosophila* protein associated with microtubules. *Cell Motil. Cytoskeleton* **63**, 301–312 (2006).
15. Lukinavičius, G. *et al.* Fluorogenic probes for live-cell imaging of the cytoskeleton. *Nature Methods* **11**, 731–733 (2014).
16. Goodwin, S. S. & Vale, R. D. Patronin regulates the microtubule network by protecting microtubule minus ends. *Cell* **143**, 263–274 (2010).
17. Euteneuer, U. & McIntosh, J. R. Polarity of midbody and phragmoplast microtubules. *J. Cell Biol.* **87**, 509–515 (1980).
18. Schiel, J. A. *et al.* Endocytic membrane fusion and buckling-induced microtubule severing mediate cell abscission. *J. Cell Sci.* **124**, 1411–1424 (2011).
19. Januschke, J., Llamazares, S., Reina, J. & Gonzalez, C. *Drosophila* neuroblasts retain the daughter centrosome. *Nature Commun.* **2**, 243 (2011).
20. Januschke, J. *et al.* Centrobin controls mother–daughter centriole asymmetry in *Drosophila* neuroblasts. *Nature Cell Biol.* **15**, 241–248 (2013).
21. Jauffred, B. *et al.* Regulation of centrosome movements by numb and the collapsin response mediator protein during *Drosophila* sensory progenitor asymmetric division. *Development* **140**, 2657–2668 (2013).
22. Ems-McClung, S. C. & Walczak, C. E. Kinesin-13s in mitosis: key players in the spatial and temporal organization of spindle microtubules. *Semin. Cell Dev. Biol.* **21**, 276–282 (2010).
23. Wang, H., Brust-Mascher, I., Civelekoglu-Scholey, G. & Scholey, J. M. Patronin mediates a switch from kinesin-13-dependent poleward flux to anaphase B spindle elongation. *J. Cell Biol.* **203**, 35–46 (2013).
24. Connell, J. W., Lindon, C., Luzio, J. P. & Reid, E. Spastin couples microtubule severing to membrane traffic in completion of cytokinesis and secretion. *Traffic* **10**, 42–56 (2009).
25. Matsuo, M. *et al.* Katanin p60 contributes to microtubule instability around the midbody and facilitates cytokinesis in rat cells. *PLoS ONE* **8**, e80392 (2013).
26. Yang, D. *et al.* Structural basis for midbody targeting of spastin by the ESCRT-III protein CHMP1B. *Nature Struct. Mol. Biol.* **15**, 1278–1286 (2008).
27. Hendershott, M. C. & Vale, R. D. Regulation of microtubule minus-end dynamics by CAMSAPs and Patronin. *Proc. Natl Acad. Sci. USA* **111**, 5860–5865 (2014).
28. Jiang, K. *et al.* Microtubule minus-end stabilization by polymerization-driven CAMSAP deposition. *Dev. Cell* **28**, 295–309 (2014).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank A. Gautreau, in whose laboratory the GBP was cloned, for allowing us to use it before his own publication. We thank A. Gautreau, R. Vale, A. Houdusse, A. Roux, A. Trushko and V. Stoppin-Mellet for discussions and advices. We thank N. Chiaruttini for the preparation of liposomes and for insights throughout the project. We thank R. Le Borgne, S. Eaton, S. De Renzis, F. Karch, W. Zhong and R. Stanewsky for flies. We thank P. Kapusta from Picoquant for help with FLIM data analysis. We thank V. Stoppin-Mellet, M. Vantard and J. Gaillard for the gift of GFP-MAP65-1. We are indebted to J. Teyssier, E. Allémann and N. Boulens for letting us use their scanning electron microscopes. E.D. is supported by a Long term Fellowship of the Human Frontier Science Program. L.H. and S.L. are supported by a Marie-Curie Intra-European Fellowship. M.G.G. is supported by Département de l'Instruction Publique of the Canton of Geneva, the SNF, the SystemsX epiPhysX program, the NCCR Chemical Biology program, the ERC (Sara and Morphogen) and the Polish–Swiss research programs.

**Author Contributions** E.D. conducted most of the biochemical experiments, imaging and image analysis and wrote the Matlab and ImageJ codes for spindle/endosome tracking and cell averaging with help from L.H. C.S. generated the anti-Klp98A antibody and characterized the four Klp98A mutants used in this study, with help from S.L. C.S., A.D. and S.L. performed additional imaging and genetic experiments. The model of endosome motility was developed jointly by E.D., M.G.G. and F.J. M.G.G. and E.D. designed the project. All authors contributed to writing the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.G.G. ([marcos.gonzalez@unige.ch](mailto:marcos.gonzalez@unige.ch)), F.J. ([julicher@pks.mpg.de](mailto:julicher@pks.mpg.de)) or E.D. ([emmanuel.derivery@unige.ch](mailto:emmanuel.derivery@unige.ch)).

## METHODS

**Generation of *Klp98A* mutants.** *Klp98A*<sup>Δ47</sup> is a null allele generated by homologous recombination with the 'ends-out' strategy<sup>29,30</sup> using a pW25 plasmid containing two homology fragments flanking the coding region of *Klp98A* as described<sup>31</sup> (2,848 bp of homology in the 5' region and 4,011 bp in the 3' region, Extended Data Fig. 1g). Upon recombination, this construct replaces the coding sequence of *Klp98A* by the *w<sup>hs</sup>* gene flanked by two *loxP* sites followed by an AttP ΦC31 site. The *w<sup>hs</sup>* gene is subsequently floxed to generate *Klp98A*<sup>Δ47</sup> which corresponds to a deletion of the *Klp98A* coding sequence (Extended Data Fig. 1h). Gene deletion was confirmed by PCR (Extended Data Fig. 1i) and sequencing.

Three zinc-finger nuclease pairs targeting *Klp98A* were designed and produced by Sigma-Aldrich (product number CSTZFNY-1KT, lot number 03041026MN). The target sequences were (cut site indicated with underlining): no. 1, CAGAGCACTGGGCATGGGCTAAGGGTGC GGAGCATCG; no. 2, CTTCGACTACTCTATTGCTCATTCGATGCGGAGGATCCG; and no. 3, CTCCTTTGCCCGCATGCGTGTGGCCAGGAGTCGGGCA.

The mRNAs corresponding to the three pairs were injected together at 40 ng μl<sup>-1</sup> in *w<sup>1118</sup>* embryos by BestGene Inc. Adults from these embryos were crossed with *w;Df(3R)BSC497/TM6c*. *Df(3R)BSC497* is a deletion spanning the *Klp98A* gene (Flybase and our own unpublished data). The relevant progeny (about 50 individuals) was then analysed by PCR using primers flanking the three cut sites and the amplicons sequenced. We found deletions only in the region corresponding to the zinc-finger pair no. 1. We studied three of them in more detail: *Klp98A*<sup>Δ6</sup>, *Klp98A*<sup>Δ7</sup> and *Klp98A*<sup>Δ8</sup>. *Klp98A*<sup>Δ6</sup> is a G to C substitution at position 500 of the coding sequence of *Klp98A* (CG5658-PA) followed by a six-nucleotide deletion. The amino acid sequence at position 167 is therefore changed from 164TGHLRVRE172 to 164TGHGA—VRE170 (see Extended Data Fig. 1j). This two amino acid deletion maps into the L8 loop of the motor domain of *Klp98A* and does not affect the stability of the protein (see Extended Data Fig. 1k) but behaves like a strong mutant in transheterozygosity with *Klp98A*<sup>Δ47</sup> (see Fig. 1h, Extended Data Fig. 3a, b, d–f). *Klp98A*<sup>Δ7</sup> is a deletion of seven nucleotides at position 502 in the coding sequence of *Klp98A*, leading to a frameshift starting at amino acid 168 and causing a stop codon after amino acid 209. *Klp98A*<sup>Δ8</sup> is a deletion of eight nucleotides at position 501, leading to a frameshift starting at amino acid 168 and causing a stop codon after amino acid 186.

Full-length *Klp98A* protein is undetectable in homozygous *Klp98A*<sup>Δ47</sup>, *Klp98A*<sup>Δ7</sup> and *Klp98A*<sup>Δ8</sup> animals (Extended Data Fig. 1a, k). In this work, transheterozygous animals (that is, *Klp98A*<sup>Δ47/Δ6</sup> and *Klp98A*<sup>Δ47/Δ8</sup>) were used in phenotypic analyses in order to avoid the effects of potential linked mutations. These transheterozygous combinations are viable and fertile. However, these mutants show Notch-dependent asymmetric cell fate assignment phenotypes when the other two systems controlling these events, that is, Numb and Neuralized, are compromised (see Fig. 1i, j and Extended Data Fig. 3).

**Fly strains.** Transgenes used in this study included *Ubi > mCherry-Pavarotti* (generated for this study), *UAS > Jupiter-mCherry* (this study), *UAS > Klp98A-mCherry* (this study), *UAS > Klp98A-GFP* (this study), *UAS-GFP-Patronin* (this study), *Asense > GFP-Pon* (this study), *Asense > mCherry-Pon* (this study), *Asense > GFP-Sara* (this study), *UAS-GBP-Pon* (this study), *UAS-GBP-mCherry-Pon* (this study), *UAS-GBP-Bazooka* (this study), *Jupiter-GFP* knock-in at the endogenous locus (ref. 14, Bloomington no. 6836), *UAS-mRFP-Pon* (ref. 32), *UAS-mRFP-Sara* (ref. 1), *Neur > Gal4* (ref. 33), *Ubi > GFP-Pavarotti* (ref. 34), *UAS > GFP-Pon* (ref. 35), *pnr > Gal4*, *phyllopod > GFP-Pon* (ref. 36), *pnr > Gal4* (Bloomington no. 3039), *UAS > DsRed* (kind gift from François Karch), *UAS > Patronin<sup>RNAi#1</sup>* (VDR no. 108927, referred to as *Patronin RNAi* in the main text), *UAS > Patronin<sup>RNAi#2</sup>* (VDR no. 27654), *UAS > Klp10A<sup>RNAi</sup>* (ref. 37, VDR no. 41534), *UAS > Klp98A<sup>RNAi</sup>* (VDR no. 40605), *UAS > Neuralized<sup>RNAi</sup>* (VDR no. 108239), *UAS > Numb<sup>RNAi</sup>* (gift from W. Zhong, ref. 38), *Df(3R)BSC497* (Bloomington no. 25001), *Klp98A*<sup>Δ47</sup> (this study), *Klp98A*<sup>Δ6</sup> (this study), *Klp98A*<sup>Δ7</sup> (this study), *Klp98A*<sup>Δ8</sup> (this study), *Numb<sup>SW</sup>* (ref. 9, gift from R. Stanewsky), *Numb<sup>2</sup>* (kind gift from Roland Le Borgne), *Numb<sup>15</sup>* (kind gift from Roland Le Borgne), *UAS > lgl3A* (ref. 39), *GFP-Rab5* knock-in at the endogenous locus (ref. 40), *YFP-Rab11* knock-in at the endogenous locus (ref. 41), *YFP-Rab7* knock-in at the endogenous locus (ref. 41) and *tub > Gal80<sup>ts</sup>* (Bloomington no. 7017). The genotypes of mutant stocks were verified by PCR and sequencing, as well as the genotypes of the F1 progeny generated for interaction studies (Fig. 1i, j and Extended Data Fig. 3). Since the Jupiter-GFP gene trap is viable, fertile and does not induce visible phenotypes in the SOP lineage, we used it as an alternative to balancers for controls in gene interaction studies (Extended Data Fig. 3). Flies co-expressing GBP-Pon and GFP-Patronin (Fig. 4 and Extended Data Figs 8 and 9) displayed occasional polarity defects reflected by loss of mRFP-Pon asymmetry (See Extended Data Fig. 8d for quantification). Cells showing such polarity defects were excluded from subsequent analysis. We used *Gal80<sup>ts</sup>* to achieve low levels of *Klp98A*-GFP expression to prevent endosome fusion (Fig. 1a, Extended Data Figs 1c–e and 2a, b).

**Detailed genotypes and temperatures.** Fig. 1a: *w<sup>1118</sup>; Neur > Gal4 UAS > mRFP-Pon tub > gal80<sup>ts</sup>/UAS > Klp98A-GFP* (25°C).

Fig. 1e–g: *w<sup>1118</sup>; Asense > GFP-Pon/Ubi > mCherry-Pavarotti* (25°C). *w<sup>1118</sup>; Asense > GFP-Pon/Ubi > mCherry-Pavarotti; Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ8</sup>* (25°C).

Fig. 1h: control: *w<sup>1118</sup>; Asense > GFP-Pon/Ubi > mCherry-Pavarotti* (25°C). *Klp98A<sup>Δ47/Δ6</sup>; w<sup>1118</sup>; Neur > Gal4 UAS > mRFP-Pon Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ6</sup>* (25°C). *Klp98A<sup>Δ47/Δ8</sup>; w<sup>1118</sup>; Asense > GFP-Pon/Ubi > mCherry-Pavarotti; Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ8</sup>* (25°C). Control RNAi: *w<sup>1118</sup>; pnr > Gal4, phyllopod > GFP-Pon/+* (29°C). *Klp98A RNAi: w<sup>1118</sup>; pnr > Gal4, phyllopod > GFP-Pon/UAS > Klp98A<sup>RNAi</sup>* (29°C).

Fig. 1i, j: *w<sup>1118</sup>* (25°C). *w<sup>1118</sup>; UAS > Neur<sup>RNAi/+</sup>; pnr > Gal4 UAS > DsRed Klp98A<sup>Δ47</sup>/TM6B* (29°C). *w<sup>1118</sup>; UAS > Neur<sup>RNAi/+</sup>; pnr > Gal4 UAS > DsRed Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ8</sup>* (29°C, sibling of fly above).

Fig. 2b–g, i–k: *w<sup>1118</sup>; Neur > Gal4, UAS > mRFP-Pon, Ubi > GFP-Pavarotti/+* (25°C).

Fig. 3a: *w<sup>1118</sup>; UAS > Jupiter-mCherry/+; Neur > Gal4, Ubi > Pavarotti-GFP/+* (25°C). *w<sup>1118</sup>; Ubi > mCherry-Pavarotti/+; Neur > Gal4, tub > Gal80<sup>ts</sup>/UAS > GFP-Patronin* (25°C). *w<sup>1118</sup>; Neur > Gal4, UAS > mRFP-Pon, Ubi > Pavarotti-GFP/+* (25°C, for SiR-Tubulin column).

Fig. 3d: *w<sup>1118</sup>; UAS > mRFP-Pon/+; Neur > Gal4, Jupiter-GFP/+* (25°C).

Fig. 3e, g, h: control: *w<sup>1118</sup>; UAS > mRFP-Pon/+; Neur > Gal4, Jupiter-GFP/+* (29°C). *Patronin RNAi: w<sup>1118</sup>; UAS > Patronin<sup>RNAi#1</sup>/UAS > mRFP-Pon; Neur > Gal4, Jupiter-GFP/+* (29°C). *Klp10A RNAi: w<sup>1118</sup>; Asense > mCherry-Pon/+; pnr > Gal4, Jupiter-GFP/UAS > Klp10A<sup>RNAi</sup>* (29°C).

Fig. 3f: *w<sup>1118</sup>; UAS > Jupiter-mCherry; Neur > Gal4, Ubi > GFP-Pavarotti/+* (25°C).

Fig. 4a, b: control: *w<sup>1118</sup>; UAS > mRFP-Pon/+; Neur > Gal4, Jupiter-GFP/+* (29°C). *Patronin RNAi: w<sup>1118</sup>; UAS > Patronin<sup>RNAi#1</sup>/UAS > mRFP-Pon; Neur > Gal4, Jupiter-GFP/+* (29°C). *Klp10A RNAi: w<sup>1118</sup>; UAS > Jupiter-mCherry/+; pnr > Gal4, phyllopod > GFP-Pon/UAS > Klp10A<sup>RNAi</sup>* (29°C).

Fig. 4c: *w<sup>1118</sup>; UAS > mRFP-Pon/+; Neur > Gal4, Jupiter-GFP/+* (25°C). *w<sup>1118</sup>; UAS > Patronin<sup>RNAi#1</sup>/UAS > mRFP-Pon; Neur > Gal4, Jupiter-GFP/+* (29°C). *w<sup>1118</sup>; Asense > mCherry-Pon/+; pnr > Gal4, Jupiter-GFP/UAS > Klp10A<sup>RNAi</sup>* (29°C). *w<sup>1118</sup>; Neur > Gal4, UAS > mRFP-Pon/UAS > GFP-Patronin* (25°C).

Fig. 4f: *w<sup>1118</sup>; Neur > Gal4, UAS > mRFP-Pon/UAS > GFP-Patronin* (25°C).

Fig. 4g: *w<sup>1118</sup>; UAS > GBP-Pon/+; Neur > Gal4, UAS > mRFP-Pon/UAS > GFP-Patronin* (25°C).

Fig. 4h: *w<sup>1118</sup>; Neur > Gal4, UAS > mRFP-Pon/UAS > GFP-Patronin* (25°C). *w<sup>1118</sup>; UAS > GBP-Pon/+; Neur > Gal4, UAS > mRFP-Pon/UAS > GFP-Patronin* (25°C).

Fig. 4i: *w<sup>1118</sup>; Neur > Gal4, UAS > mRFP-Pon/UAS > GFP-Patronin* (25°C). *w<sup>1118</sup>; UAS > GBP-Pon/+; Neur > Gal4, UAS > mRFP-Pon/UAS > GFP-Patronin* (25°C). *w<sup>1118</sup>; UAS > mRFP-Pon/+; Neur > Gal4, Jupiter-GFP/+* (25°C). *w<sup>1118</sup>; UAS > Patronin<sup>RNAi#1</sup>/UAS > mRFP-Pon; Neur > Gal4, Jupiter-GFP/+* (29°C). *w<sup>1118</sup>; Asense > mCherry-Pon/+; pnr > Gal4, Jupiter-GFP/UAS > Klp10A<sup>RNAi</sup>* (29°C).

Extended Data Fig. 1a: *w<sup>1118</sup>* (25°C). *w<sup>1118</sup>; Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ47</sup>* (25°C). *w<sup>1118</sup>; Klp98A<sup>Δ8</sup>/Klp98A<sup>Δ8</sup>* (25°C).

Extended Data Fig. 1b: *w<sup>1118</sup>; Neur > Gal4, UAS > mRFP-Pon/+* (25°C). *w<sup>1118</sup>; Asense > GFP-Pon/Ubi > mCherry-Pavarotti; Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ8</sup>* (25°C).

Extended Data Fig. 1c: iDelta20 endogenous *Klp98A*: *w<sup>1118</sup>; Neur > Gal4 UAS > mRFP-Pon* (25°C). GFP-Sara endogenous *Klp98A*: *w<sup>1118</sup>; Asense > GFP-Sara/+; Neur > Gal4 UAS > mRFP-Pon/+* (29°C). GFP-Sara *Klp98A*-mCherry: *w<sup>1118</sup>; Asense > GFP-Sara/+; Neur > Gal4 tub > gal80<sup>ts</sup> UAS > Klp98A-mCherry/+* (29°C). GFP-Rab5 *Klp98A*-mCherry: *w<sup>1118</sup>; GFP-Rab5/+; Neur > Gal4 tub > gal80<sup>ts</sup> UAS > Klp98A-mCherry/+* (25°C). YFP-Rab7 *Klp98A*-mCherry: *w<sup>1118</sup>; Neur > Gal4 tub > gal80<sup>ts</sup> UAS > Klp98A-mCherry/YFP-Rab7* (25°C). YFP-Rab11 *Klp98A*-mCherry: *w<sup>1118</sup>; Neur > Gal4 tub > gal80<sup>ts</sup> UAS > Klp98A-mCherry/YFP-Rab11* (25°C).

Extended Data Fig. 1d: GFP-Sara *Klp98A*-mCherry: *w<sup>1118</sup>; Asense > GFP-Sara/+; Neur > Gal4 tub > gal80<sup>ts</sup> UAS > Klp98A-mCherry/+* (29°C). GFP-Rab5 *Klp98A*-mCherry: *w<sup>1118</sup>; GFP-Rab5/+; Neur > Gal4 tub > gal80<sup>ts</sup> UAS > Klp98A-mCherry/+* (25°C). YFP-Rab7 *Klp98A*-mCherry: *w<sup>1118</sup>; Neur > Gal4 tub > gal80<sup>ts</sup> UAS > Klp98A-mCherry/YFP-Rab7* (25°C).

Extended Data Fig. 1e: control: *w<sup>1118</sup>; Asense > GFP-Sara/+; Neur > Gal4 UAS > mRFP-Pon/+* (29°C). *Klp98A*-mCherry: *w<sup>1118</sup>; Asense > GFP-Sara/+; Neur > Gal4 tub > gal80<sup>ts</sup> UAS > Klp98A-mCherry/+* (29°C).

Extended Data Fig. 1i: *w<sup>1118</sup>* (25°C). *w<sup>1118</sup>; Df(3R)BSC497/Klp98A<sup>Δ47</sup>* (25°C).

Extended Data Fig. 1k: *w<sup>1118</sup>* (25°C). *w<sup>1118</sup>; Klp98A<sup>Δ6</sup>/Klp98A<sup>Δ6</sup>* (25°C). *w<sup>1118</sup>; Klp98A<sup>Δ7</sup>/Klp98A<sup>Δ7</sup>* (25°C). *w<sup>1118</sup>; Klp98A<sup>Δ8</sup>/Klp98A<sup>Δ8</sup>* (25°C).

Extended Data Fig. 2a, b: control: *w<sup>1118</sup>; Asense > GFP-Sara/+; Neur > Gal4 UAS > mRFP-Pon/+* (29°C). *Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ6</sup>; w<sup>1118</sup>; Asense > GFP-Sara Asense-mCherry-Pon/+; Klp98A<sup>Δ7</sup>/Klp98A<sup>Δ6</sup>* (29°C). *Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ8</sup>; w<sup>1118</sup>; Asense >*



GFP-Sara Asense-mCherry-Pon/+;Klp98A<sup>Δ7</sup>/Klp98A<sup>Δ8</sup> (29°C). Patronin RNAi: w<sup>1118</sup>;Asense > GFP-Sara/UAS > Patronin<sup>RNAi#1</sup>;Neur > Gal4 UAS > mRFP-Pon/+ (29°C). Klp10A RNAi: w<sup>1118</sup>;Asense > GFP-Sara/+;pnr > Gal4 UAS > DsRed/UAS > Klp10A<sup>RNAi</sup> (29°C). Klp98A-mCherry: w<sup>1118</sup>;Asense > GFP-Sara/+;Neur > Gal4 tub > gal80<sup>ts</sup> UAS > Klp98A-mCherry/+ (29°C). GFP-Patronin + GBP-Pon: w<sup>1118</sup>;UAS > GBP-Pon/tub > gal80<sup>ts</sup>;UAS > GFP-Patronin/Neur > Gal4 UAS > mRFP-Sara (25°C).

Extended Data Fig. 2c: w<sup>1118</sup>;Asense > GFP-Sara/+;Neur > Gal4 UAS > mRFP-Pon/+ (29°C).

Extended Data Fig. 2d: w<sup>1118</sup>;Asense > GFP-Sara Asense-mCherry-Pon/+;Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ8</sup> (29°C).

Extended Data Fig. 2e: w<sup>1118</sup>;Asense > GFP-Sara/UAS > Patronin<sup>RNAi#1</sup>;Neur > Gal4 UAS > mRFP-Pon/+ (29°C).

Extended Data Fig. 2f: w<sup>1118</sup>;Asense > GFP-Sara/+;pnr > Gal4 UAS > DsRed/UAS > Klp10A<sup>RNAi</sup> (29°C).

Extended Data Fig. 2g: w<sup>1118</sup>;Asense > GFP-Sara/+;Neur > Gal4 tub > gal80<sup>ts</sup> UAS > Klp98A-mCherry/+ (29°C).

Extended Data Fig. 2h: w<sup>1118</sup>;UAS > GBP-Pon/tub > gal80<sup>ts</sup>;UAS > GFP-Patronin/Neur > Gal4 UAS > mRFP-Sara (25°C).

Extended Data Fig. 2i: w<sup>1118</sup>;Asense > GFP-Pon/Ubi > mCherry-Pavarotti (25°C). w<sup>1118</sup>;Asense > GFP-Pon/Ubi > mCherry-Pavarotti;Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ8</sup> (25°C).

Extended Data Fig. 3a, b: w<sup>1118</sup> (25°C). w<sup>1118</sup>;UAS > Neur<sup>RNAi</sup>/+;pnr > Gal4 UAS > DsRed Klp98A<sup>Δ47</sup>/TM6B (29°C). w<sup>1118</sup>;UAS > Neur<sup>RNAi</sup>/+;pnr > Gal4 UAS > DsRed Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ6</sup> (29°C; sibling of fly above).

Extended Data Fig. 3c: Klp98A<sup>Δ47/+</sup>; w<sup>1118</sup>;UAS > Neur<sup>RNAi</sup>/+;pnr > Gal4 UAS > DsRed Klp98A<sup>Δ47</sup>/TM6B (29°C; outside the pnr expression region). pnr > Neur<sup>RNAi</sup> Klp98A<sup>Δ47/+</sup>; w<sup>1118</sup>;UAS > Neur<sup>RNAi</sup>/+;pnr > Gal4 UAS > DsRed Klp98A<sup>Δ47</sup>/TM6B (29°C; inside the pnr expression region). pnr > Neur<sup>RNAi</sup> Klp98A<sup>Δ47/+</sup>; w<sup>1118</sup>;UAS > Neur<sup>RNAi</sup>/+;pnr > Gal4 UAS > DsRed Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ8</sup> (29°C; inside the pnr expression region; sibling of fly above).

Extended Data Fig. 3d: w<sup>1118</sup> (25°C). numb<sup>2/SW</sup> Klp98A<sup>Δ47</sup>/Jupiter-GFP: w<sup>1118</sup>(+);Numb<sup>2</sup>/Numb<sup>SW</sup>;Jupiter-GFP/Klp98A<sup>Δ47</sup> (25°C). numb<sup>2/SW</sup> Klp98A<sup>Δ47/Δ8</sup>: w<sup>1118</sup>(+);Numb<sup>2</sup>/Numb<sup>SW</sup>;Klp98A<sup>Δ8</sup>/Klp98A<sup>Δ47</sup> (25°C; sibling of the fly above). pnr > numb<sup>RNAi</sup> Klp98A<sup>Δ47/+</sup>; w<sup>1118</sup>;pnr > Gal4 UAS > DsRed Klp98A<sup>Δ47</sup>/UAS > numb<sup>RNAi</sup> (29°C). pnr > numb<sup>RNAi</sup> Klp98A<sup>Δ47/Δ8</sup>; w<sup>1118</sup>;pnr > Gal4 UAS > DsRed Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ8</sup> UAS > numb<sup>RNAi</sup> (29°C).

Extended Data Fig. 3e: numb<sup>2/SW</sup> Klp98A<sup>Δ47</sup>/Jupiter-GFP: w<sup>1118</sup>(+);Numb<sup>2</sup>/Numb<sup>SW</sup>;Jupiter-GFP/Klp98A<sup>Δ47</sup> (25°C). numb<sup>2/SW</sup> Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ6</sup>: w<sup>1118</sup>(+);Numb<sup>2</sup>/Numb<sup>SW</sup>;Klp98A<sup>Δ6</sup>/Klp98A<sup>Δ47</sup> (25°C; sibling of the fly above). numb<sup>2/SW</sup> Klp98A<sup>Δ47</sup>/Jupiter-GFP: w<sup>1118</sup>(+);Numb<sup>2</sup>/Numb<sup>SW</sup>;Jupiter-GFP/Klp98A<sup>Δ47</sup> (25°C). numb<sup>2/SW</sup> Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ8</sup>: w<sup>1118</sup>(+);Numb<sup>2</sup>/Numb<sup>SW</sup>;Klp98A<sup>Δ8</sup>/Klp98A<sup>Δ47</sup> (25°C; sibling of the fly above). numb<sup>15/SW</sup> Klp98A<sup>Δ47/+</sup>; w<sup>1118</sup>(+);Numb<sup>15</sup>/Numb<sup>SW</sup>;Klp98A<sup>Δ47</sup>/TM6B (25°C) or w<sup>1118</sup>(+);Numb<sup>15</sup>/Numb<sup>SW</sup>;Klp98A<sup>Δ6</sup>/TM6B (25°C). numb<sup>15/SW</sup> Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ6</sup>; w<sup>1118</sup>(+);Numb<sup>2</sup>/Numb<sup>SW</sup>;Klp98A<sup>Δ6</sup>/Klp98A<sup>Δ47</sup> (25°C; sibling of the fly above).

Extended Data Fig. 3f: pnr > numb<sup>RNAi</sup> Klp98A<sup>Δ47/+</sup>; w<sup>1118</sup>;pnr > Gal4 UAS > DsRed Klp98A<sup>Δ47</sup>/UAS > numb<sup>RNAi</sup> (29°C). pnr > numb<sup>RNAi</sup> Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ6</sup>: w<sup>1118</sup>;pnr > Gal4 UAS > DsRed Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ6</sup> UAS > numb<sup>RNAi</sup> (29°C). pnr > numb<sup>RNAi</sup> Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ8</sup>: w<sup>1118</sup>;pnr > Gal4 UAS > DsRed Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ8</sup> UAS > numb<sup>RNAi</sup> (29°C).

Extended Data Fig. 4a–e: w<sup>1118</sup>;Asense > GFP-Pon/Ubi > mCherry-Pavarotti (25°C). w<sup>1118</sup>;Asense > GFP-Pon/Ubi > mCherry-Pavarotti;Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ8</sup> (25°C).

Extended Data Fig. 4f–n: w<sup>1118</sup>;Neur > Gal4, UAS > mRFP-Pon, Ubi > GFP-Pavarotti/+ (25°C).

Extended Data Fig. 4o–v: w<sup>1118</sup>;Neur > Gal4, UAS > mRFP-Pon, Ubi > GFP-Pavarotti/+ (25°C).

Extended Data Fig. 4w: w<sup>1118</sup>;Neur > Gal4, UAS > mRFP-Pon, Ubi > GFP-Pavarotti/+ (25°C). w<sup>1118</sup>;UAS > mRFP-Pon/+;Neur > Gal4, Jupiter-GFP/+ (29°C).

Extended Data Fig. 4x: control: w<sup>1118</sup>;Neur > Gal4, UAS > mRFP-Pon, Ubi > GFP-Pavarotti/+ (25°C). w<sup>1118</sup>;UAS > mRFP-Pon/+;Neur > Gal4, Jupiter-GFP/+ (29°C). Patronin RNAi#1: w<sup>1118</sup>;UAS > Patronin<sup>RNAi#1</sup>/+;Neur > Gal4, UAS > mRFP-Pon, Ubi > GFP-Pavarotti/+ (29°C) and w<sup>1118</sup>;UAS > Patronin<sup>RNAi#1</sup>/UAS > mRFP-Pon;Neur > Gal4, Jupiter-GFP/+ (29°C). Klp10A RNAi: w<sup>1118</sup>;Asense > mCherry-Pon/+;pnr > Gal4, Jupiter-GFP/UAS > Klp10A<sup>RNAi</sup> (29°C). Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ8</sup>: w<sup>1118</sup>;Asense > GFP-Pon/Ubi > mCherry-Pavarotti; Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ8</sup> (25°C).

Extended Data Fig. 4y: w<sup>1118</sup>;Neur > Gal4, UAS > mRFP-Pon, Ubi > GFP-Pavarotti/+ (25°C).

Extended Data Fig. 5a–b: w<sup>1118</sup>;UAS > Jupiter-mCherry/+;Neur > Gal4, Ubi > GFP-Pavarotti/+ (25°C).

Extended Data Fig. 5c: w<sup>1118</sup>;UAS > Jupiter-mCherry/+;Neur > Gal4, UAS > mRFP-Pon, Ubi > Pavarotti-GFP/+ (25°C).

Extended Data Fig. 5d–i: w<sup>1118</sup>;Neur > Gal4, UAS > mRFP-Pon/+ (25°C).

Extended Data Fig. 5j, k: control: w<sup>1118</sup>;UAS > mRFP-Pon/+;Neur > Gal4, Jupiter-GFP/+ (25°C). Lgl3A: w<sup>1118</sup>;UAS > Lgl3A/UAS > mRFP-Pon;Neur > Gal4, Jupiter-GFP/+ (25°C).

Extended Data Fig. 5l, n: Jupiter-GFP: w<sup>1118</sup>;UAS > mRFP-Pon/+;Neur > Gal4, Jupiter-GFP/+ (25°C). GFP-Patronin and SiR-tubulin: w<sup>1118</sup>;Neur > Gal4, UAS > mRFP-Pon/UAS > GFP-Patronin (25°C).

Extended Data Fig. 5m, n: Jupiter-GFP: w<sup>1118</sup>;UAS > mRFP-Pon/+;Neur > Gal4, Jupiter-GFP/+ (25°C). GFP-Patronin: w<sup>1118</sup>;Neur > Gal4, UAS > mRFP-Pon/UAS > GFP-Patronin (25°C). SiR-Tubulin: w<sup>1118</sup>;Neur > Gal4, UAS > mRFP-Pon, Ubi > GFP-Pavarotti/+ (25°C).

Extended Data Fig. 5o, p: control: w<sup>1118</sup>;UAS > mRFP-Pon/+;Neur > Gal4, Jupiter-GFP/+ (29°C). Patronin RNAi: w<sup>1118</sup>;UAS > Patronin<sup>RNAi#1</sup>/UAS > mRFP-Pon;Neur > Gal4, Jupiter-GFP/+ (29°C). Klp10A RNAi: w<sup>1118</sup>;Asense > mCherry-Pon/+;pnr > Gal4, Jupiter-GFP/UAS > Klp10A<sup>RNAi</sup> (29°C).

Extended Data Fig. 5q, s: EB1-GFP and SiR-tubulin: w<sup>1118</sup>;Ubi > mCherry-Pavarotti Neur > Gal4 tub > gal80<sup>ts</sup>/UAS > EB1-GFP (25°C). Jupiter-GFP: w<sup>1118</sup>;UAS > mRFP-Pon/+;Neur > Gal4, Jupiter-GFP/+ (25°C). GFP-Patronin: w<sup>1118</sup>;Neur > Gal4, UAS > mRFP-Pon/UAS > GFP-Patronin (25°C).

Extended Data Fig. 5r, s: EB1-GFP: w<sup>1118</sup>;Ubi > mCherry-Pavarotti Neur > Gal4 tub > gal80<sup>ts</sup>/UAS > EB1-GFP (25°C). Jupiter-GFP: w<sup>1118</sup>;UAS > mRFP-Pon/+;Neur > Gal4, Jupiter-GFP/+ (25°C). GFP-Patronin and SiR-tubulin: w<sup>1118</sup>;Neur > Gal4, UAS > mRFP-Pon/UAS > GFP-Patronin (25°C).

Extended Data Fig. 6b, c: control: w<sup>1118</sup>;UAS > mRFP-Pon/+;Neur > Gal4, Jupiter-GFP/+ (29°C). Patronin RNAi#1: w<sup>1118</sup>;UAS > Patronin<sup>RNAi#1</sup>/UAS > mRFP-Pon;Neur > Gal4, Jupiter-GFP/+ (29°C).

Extended Data Fig. 6d–f: control: w<sup>1118</sup>;UAS-mRFP-Pon/+;Neur > Gal4, Jupiter-GFP/+ (29°C). Patronin RNAi#1: w<sup>1118</sup>;UAS > Patronin<sup>RNAi#1</sup>/UAS > mRFP-Pon;Neur > Gal4, Jupiter-GFP/+ (29°C). Patronin RNAi#2: w<sup>1118</sup>;UAS > Patronin<sup>RNAi#2</sup>/UAS > mRFP-Pon;Neur > Gal4, Jupiter-GFP/+ (29°C).

Extended Data Fig. 6g, h: GFP-Patronin: w<sup>1118</sup>;Neur > Gal4, UAS > mRFP-Pon/UAS > GFP-Patronin (25°C). Patronin RNAi#1 + GFP-Patronin: w<sup>1118</sup>;UAS > Patronin<sup>RNAi#1</sup>/+;Neur > Gal4, UAS > mRFP-Pon, UAS > GFP-Patronin/+ (29°C).

Extended Data Fig. 6i, j: control: w<sup>1118</sup>;UAS > mRFP-Pon/+;Neur > Gal4, Jupiter-GFP/+ (29°C). Patronin RNAi#1: w<sup>1118</sup>;UAS > Patronin<sup>RNAi#1</sup>/UAS > mRFP-Pon;Neur > Gal4, Jupiter-GFP/+ (29°C). Klp10A RNAi: w<sup>1118</sup>;Asense > mCherry-Pon/+;pnr > Gal4, Jupiter-GFP/UAS > Klp10A<sup>RNAi</sup> (29°C). Klp10A RNAi + Patronin RNAi#1: w<sup>1118</sup>;UAS > Patronin<sup>RNAi#1</sup>/UAS > mRFP-Pon;Neur > Gal4, Jupiter-GFP/UAS > Klp10A<sup>RNAi</sup> (29°C).

Extended Data Fig. 7a–e: w<sup>1118</sup>;UAS > Patronin<sup>RNAi#1</sup>/+;Neur > Gal4, UAS > mRFP-Pon, Ubi > GFP-Pavarotti/+ (29°C).

Extended Data Fig. 7f: control: w<sup>1118</sup>;Asense > GFP-Pon/Ubi > mCherry-Pavarotti (25°C). Patronin RNAi: w<sup>1118</sup>;UAS > Patronin<sup>RNAi#1</sup>/UAS > mRFP-Pon;Neur > Gal4, Jupiter-GFP/+ (29°C). Klp10A RNAi: w<sup>1118</sup>;Asense > mCherry-Pon/+;pnr > Gal4, Jupiter-GFP/UAS > Klp10A<sup>RNAi</sup> (29°C).

Extended Data Fig. 7g, h: control: w<sup>1118</sup>;Neur > Gal4, UAS > mRFP-Pon/+ (29°C). Patronin RNAi#1: w<sup>1118</sup>;UAS > Patronin<sup>RNAi#1</sup>/+;Neur > Gal4, UAS > mRFP-Pon/+ (29°C). Patronin RNAi#1 + GFP-Patronin: w<sup>1118</sup>;UAS > Patronin<sup>RNAi#1</sup>/+;Neur > Gal4, UAS > mRFP-Pon, UAS > GFP-Patronin/+ (29°C).

Extended Data Fig. 7i, j: control: w<sup>1118</sup>;pnr > Gal4, phyllopod > GFP-Pon/+ (29°C). Patronin RNAi#2: w<sup>1118</sup>;UAS > Patronin RNAi#2/+;pnr > Gal4, phyllopod > GFP-Pon/+ (29°C).

Extended Data Fig. 7k, l: control: w<sup>1118</sup>;UAS > mRFP-Pon/+;Neur > Gal4, Jupiter-GFP/+ (29°C). Patronin RNAi#1: w<sup>1118</sup>;UAS > Patronin<sup>RNAi#1</sup>/UAS > mRFP-Pon;Neur > Gal4, Jupiter-GFP/+ (29°C). Klp10A RNAi: w<sup>1118</sup>;UAS > Jupiter-mCherry/+;pnr > Gal4, phyllopod > GFP-Pon/UAS > Klp10A<sup>RNAi</sup> (29°C). Klp10A RNAi + Patronin RNAi#1: w<sup>1118</sup>;UAS > Patronin<sup>RNAi#1</sup>/UAS > mRFP-Pon;Neur > Gal4, Jupiter-GFP/UAS > Klp10A<sup>RNAi</sup> (29°C).

Extended Data Fig. 8c: w<sup>1118</sup>;Neur > Gal4, UAS > mRFP-Pon/UAS > GFP-Patronin (25°C). w<sup>1118</sup>;UAS > GBP-Pon/+;Neur > Gal4, UAS > mRFP-Pon/UAS > GFP-Patronin (25°C).

Extended Data Fig. 8d: control: w<sup>1118</sup>;Neur > Gal4, UAS-mRFP > Pon/+ (25°C). GFP-Patronin: w<sup>1118</sup>;Neur > Gal4, UAS-mRFP > Pon/UAS > GFP-Patronin (25°C). GBP-Pon: w<sup>1118</sup>;UAS > GBP-Pon/+;Neur > Gal4, UAS > mRFP-Pon/+ (25°C). GBP-Pon + GFP: w<sup>1118</sup>;UAS > GBP-Pon/UAS > GFP;Neur < Gal4, UAS > mRFP-Pon/+ (25°C). GBP-Pon + GFP-Patronin: w<sup>1118</sup>;UAS > GBP-Pon/+;Neur > Gal4, UAS > mRFP-Pon/UAS > GFP-Patronin (25°C).



Extended Data Fig. 8e: GFP-Patronin:  $w^{1118};\text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon/} \text{UAS} > \text{GFP-Patronin}$  (25°C). GBP-Pon+ GFP-Patronin:  $w^{1118};\text{UAS} > \text{GBP-Pon/+}; \text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon/} \text{UAS} > \text{GFP-Patronin}$  (25°C).

Extended Data Fig. 8f, g: control:  $w^{1118};\text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon/+}$  (25°C). GBP-Pon+ GFP-Patronin:  $w^{1118};\text{UAS} > \text{GBP-Pon/+}; \text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon/} \text{UAS} > \text{GFP-Patronin}$  (25°C).

Extended Data Fig. 8h: control:  $w^{1118};\text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon/+}$  (25°C). GFP-Patronin:  $w^{1118};\text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon/} \text{UAS} > \text{GFP-Patronin}$  (25°C). GBP-Pon:  $w^{1118};\text{UAS} > \text{GBP-Pon/+}; \text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon/+}$  (25°C). GBP-Pon + GFP:  $w^{1118};\text{UAS} > \text{GBP-Pon/} \text{UAS} > \text{GFP}; \text{Neur-Gal4}, \text{UAS} > \text{mRFP-Pon/+}$  (25°C). GBP-Pon + GFP-Patronin:  $w^{1118};\text{UAS} > \text{GBP-Pon/+}; \text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon/} \text{UAS} > \text{GFP-Patronin}$  (25°C).

Extended Data Fig. 9a:  $w^{1118};\text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon/} \text{UAS-GFP-Patronin}$  (25°C).  $w^{1118};\text{Neur} > \text{Gal4}, \text{UAS} > \text{GBP-mCherry-Pon/} \text{UAS-GFP-Patronin}$  (25°C).

Extended Data Fig. 9c, d:  $w^{1118};\text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon/} \text{UAS} > \text{GFP-Patronin}$  (25°C).  $w^{1118};\text{UAS} > \text{GBP-Pon/+}; \text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon/} \text{UAS} > \text{GFP-Patronin}$  (25°C).

Extended Data Fig. 9e:  $w^{1118};\text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon/} \text{UAS} > \text{GFP-Patronin}$  (25°C).

Extended Data Fig. 9f:  $w^{1118};\text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon/} \text{UAS} > \text{GFP-Patronin}$  (25°C).  $w^{1118};\text{UAS} > \text{GBP-Bazooka/+}; \text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon}, \text{tub} > \text{gal80}^{\Delta\text{S}}/\text{UAS} > \text{GFP-Patronin}$  (25°C).

Extended Data Fig. 9g, h:  $w^{1118};\text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon/} \text{UAS} > \text{GFP-Patronin}$  (25°C).  $w^{1118};\text{UAS} > \text{GBP-Pon/+}; \text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon/} \text{UAS} > \text{GFP-Patronin}$  (25°C).  $w^{1118};\text{UAS} > \text{GBP-Bazooka/+}; \text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon}, \text{tub} > \text{gal80}^{\Delta\text{S}}/\text{UAS} > \text{GFP-Patronin}$  (25°C).

Extended Data Fig. 10b–f:  $w^{1118};\text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon}, \text{Ubi} > \text{GFP-Pavarotti/+}$  (25°C).

Extended Data Fig. 10g:  $w^{1118};\text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon/} \text{UAS} > \text{GFP-Patronin}$  (25°C).  $w^{1118};\text{UAS} > \text{GBP-Pon/+}; \text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon/} \text{UAS} > \text{GFP-Patronin}$  (25°C).

Extended Data Fig. 10h, i:  $w^{1118};\text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon/} \text{UAS} > \text{GFP-Patronin}$  (25°C).  $w^{1118};\text{UAS} > \text{GBP-Pon/+}; \text{Neur} > \text{Gal4}, \text{UAS} > \text{mRFP-Pon/} \text{UAS} > \text{GFP-Patronin}$  (25°C).  $w^{1118};\text{UAS-mRFP-Pon/+}; \text{Neur} > \text{Gal4}, \text{Jupiter-GFP/+}$  (25°C).  $w^{1118};\text{UAS} > \text{Patronin}^{\text{RNAi\#1}}/\text{UAS} > \text{mRFP-Pon}; \text{Neur} > \text{Gal4}, \text{Jupiter-GFP/+}$  (29°C).  $w^{1118};\text{Asease} > \text{mCherry-Pon/+}; \text{pnr} > \text{Gal4}, \text{Jupiter-GFP/} \text{UAS} > \text{Klp10A}^{\text{RNAi}}$  (29°C).

Most of these genotypes correspond to the F1 of crosses performed at 25°C. Embryos were laid at 25°C. Larvae were then shifted to 16°C until puparium formation and 16h before SOP imaging they were shifted to 25°C or 29°C, as indicated. In Extended Data Fig. 1a, k, larvae homozygous for the *Klp98A* mutants were used for western blot analysis instead of animals deriving from an outcross.

**Plasmids.** All the open reading frames (ORFs) cloned by PCR for this study were flanked by FseI and AscI sites for convenient shuttling between compatible plasmids. eGFP was amplified from pEGFP C1 (Clontech). *Pavarotti* (CG1258-PA), *Klp98A* (CG5658-PA), *Bazooka* (CG5055-PA) and *Patronin* (CG33130) were amplified from cDNAs prepared from adult  $w^{1118}$  flies total RNA extracted in TRIzol (Life Technologies), followed by reverse transcription (Super Script II kit, Life Technologies). The *Patronin* cDNA that we cloned encodes a splicing isoform slightly smaller than previously reported *Patronin* cDNAs<sup>16,23</sup> and has been deposited in the NCBI database (BankIt1865736 Patronin KT953618). The Pon localization domain (corresponding to amino acids 474–670 of Pon<sup>35</sup>) was similarly cloned from cDNA from  $w^{1118}$  flies. In various transgenes in this work (driven by UAS or Ase promoters), this Pon localization domain is referred to as 'Pon' for simplicity. Sara was subcloned from pUAST-GFP-Sara<sup>42</sup>. For antibody production, we also cloned smaller fragments of Patronin (corresponding to amino acids 1,039–1,384, named Patronin-Cter thereafter) and Klp98A (corresponding to amino acids 401–1,265, named Klp98A-Cter thereafter). Jupiter-mCherry was generated by cloning Jupiter-GFP from cDNA prepared from Jupiter-GFP flies<sup>14</sup> and by replacing the GFP (in the middle of the gene) by mCherry using site-directed mutagenesis. We also cloned the GFP-binding peptide (GBP), or so called GFP nanobody, a llama VHH single chain antibody against GFP<sup>43</sup> either for protein production (His-GBP) or for expression of fusion proteins in the fly (GBP-Pon and GBP-Bazooka, see below).

Klp98A-GFP-PC: for stable expression of Klp98A in S2 cells, the Klp98A ORF described above was subcloned into a modified pMT vector (Life Technologies), to which a Puromycin selection gene (amplified from the pCoPuro plasmid<sup>44</sup>) and a C-terminal tag (eGFP followed by PC, the Protein C epitope tag: EDQVDPRLIDG) were added.

GST-Klp98A-Cter, and GST-Patronin-Cter: for expression of GST-Klp98A-Cter and GST-Patronin-Cter in bacteria, the ORFs described above were subcloned into a modified pGEX vector<sup>45</sup>.

His-Klp98A-Cter, His-eGFP, and His-GBP: for expression of (His)<sub>6</sub>-tagged Klp98A-Cter, GFP and GBP, these ORFs were subcloned into a modified pET28b vector, which tags the ORF at its N terminus with a (His)<sub>6</sub> tag.

UAS > GFP-Patronin, UAS > Jupiter-mCherry, UAS > GBP-Pon, UAS > GBP-mCherry-Pon UAS > GBP-Bazooka, UAS > Klp98A-mCherry, UAS > Klp98A-GFP: for expression in flies with the UAS/Gal4 system, the Patronin, Bazooka, Klp98A, Pon localization domain and Jupiter-mCherry ORFs described above were subcloned into modified pUAST4 vectors tagging the ORF with either N-terminal PC-eGFP (for Patronin, referred to as GFP-Patronin for simplicity), C-terminal mCherry-PC (for Klp98A), C-terminal eGFP (for Klp98A), N-terminal GBP (for the Pon localization domain and Bazooka), N-terminal GBP-mCherry (GBP followed by mCherry separated by a GGG linker, for the Pon localization domain) or leaving it untagged (for Jupiter-mCherry). N-terminal GFP tagging of Patronin has been previously shown to be functional<sup>16</sup>, as well as N-terminal tagging of Bazooka<sup>46</sup>.

Ubi > mCherry-Pavarotti: for ubiquitous expression of mCherry-Pavarotti, Pavarotti was subcloned into a modified pUbi vector allowing the expression of mCherry-Pavarotti under the ubiquitin promoter.

Ase > GFP-Pon, Ase > GFP-Sara and Ase > mCherry-Pon: for specific expression in SOPs independently of the UAS/Gal4 system, the Pon localization domain and Sara were subcloned into the pSense GFP vector, which was created by inserting a 1,943-bp fragment upstream of the start codon of the *Ase* gene (amplified from  $w^{1118}$  flies genomic DNA) into the Green Pelican GFP plasmid (*Drosophila* Genomics Resource Center), which results in tagging the Pon localization domain (or Sara) with an N-Terminal GFP (Ase > GFP-Pon). Alternatively, the GFP was exchanged by quick-change PCR into mCherry to generate the pSense mCherry vector, in which the Pon localization domain was subcloned.

Injection of plasmids into *Drosophila* embryos to generate transgenics was performed by BestGene Inc.

**SDS-PAGE and western blot.** SDS-PAGE was performed using NuPAGE 4–12% Bis-Tris gels (Life Technologies) according to the manufacturer's instructions. Colloidal Coomassie blue (Life Technologies) was used for total protein staining of gels. Gels were transferred on nitrocellulose membranes using iBLOT (Life Technologies) according to the manufacturer's instructions. For western blot, we used all primary antibodies at 1 µg ml<sup>-1</sup> in TBS, 0.2% BSA, 1 mM CaCl<sub>2</sub>, 0.02% Thymersal O/N at 4°C. Western blots were revealed using HRP coupled antibodies (Jackson ImmunoResearch 1:10,000 dilution), Western Bright Quantum (Advanta) or SuperSignal West Pico (Pierce) chemiluminescence reagents and a Vilber Lourmat Fusion imager. Alternatively (Extended Data Fig. 6a), western blots were performed with fluorescently labelled anti-tubulin antibodies and imaged with an Ettan DIGE Imager (GE Healthcare). For gel source data, see Supplementary Fig. 1.

**Fly and S2 cell total extracts.** For total fly extracts (Extended Data Fig. 1a, k), dissected brains, imaginal discs and salivary glands of second instar larvae were squashed into 500 µl of lysis buffer (25 mM NaF, 1 mM Na<sub>3</sub>VO<sub>4</sub>, 50 mM Tris pH 7.5, 1.5 mM MgCl<sub>2</sub>, 125 mM NaCl, 0.2% IGEPAL, 5% glycerol, 1 mM DTT and protease inhibitor cocktail (benzamide (1 mM, Applchem), chymostatin (40 µg ml<sup>-1</sup>, Applchem), antipain (40 µg ml<sup>-1</sup>, Applchem), leupeptin (1 µM, Applchem), pefabloc (1 mM) and PMSF (0.5 mM)). The extract was incubated 40 min at 4°C with rocking, then cellular debris were cleared by centrifugation at 16,000g for 10 min at 4°C. Extracts were then diluted in LDS sample buffer (Life Technologies) enriched with 2.5% β-mercaptoethanol and analysed by SDS-PAGE and western blot as above.

For RNAi-treated S2 total cell extracts (Extended Data Fig. 6a), *Drosophila* S2 cells (UCSF, mycoplasma-free judged by DAPI staining) were cultured and incubated with 5 µg dsRNA for 4 days as previously described<sup>47</sup>. This dsRNA sequence corresponds to the sequence in the UAS > *Patronin*<sup>RNAi\#1</sup> fly stock (VDR no. 108927). Cells were washed in XB (20 mM HEPES, 150 mM KCl, pH 7.7), resuspended in LDS sample buffer, boiled for 2 min, then treated with Benzonase (30 units µl<sup>-1</sup>, Sigma) and analysed by SDS-PAGE and western blot as above.

**Protein purification.** Unless stated otherwise, reagents were from Sigma. All purification steps were performed at 4°C. Protein concentrations were determined spectrophotometrically using absorbance at 280 nm or after SDS-PAGE using purified BSA as a standard, followed by quantifications by densitometry using ImageJ (<http://imagej.nih.gov/ij/>).

GST- and His-tagged Klp98A-Cter were expressed in *E. coli* BL21 Rosetta 2 (Stratagene) by induction with 0.5 mM IPTG in Terrific Broth medium (Sigma) at 23°C. Bacteria expressing GST-Klp98A-Cter were lysed enzymatically using 0.7 mg ml<sup>-1</sup> lysosyme and 10 µg ml<sup>-1</sup> DNase I (Roche) in lysis buffer (50 mM Tris, 150 mM NaCl, 1% Triton X-100, 1 mM DTT, 5% Glycerol, pH 7.6) enriched with protease inhibitors (Roche Mini) for 1 h at 4°C with rocking. After clarification (12,000 r.p.m., Beckman JA 25.5), lysate was incubated with glutathione sepharose resin (glutathione sepharose 4B, Amersham) for 2 h at 4°C and washed extensively

in 50 mM Tris, 2 mM  $\beta$ -mercaptoethanol, 100 mM NaCl, 5 mM  $MgCl_2$  pH 7.5. Glutathione-sepharose-bound GST-Klp98A-Cter was then cleaved on column by an overnight incubation at 4°C with 40  $\mu$ g of TEV protease per mg of fusion protein. Klp98A-Cter was subsequently dialysed against PBS, concentrated to 1 mg ml<sup>-1</sup> by ultrafiltration (Amicon Ultra-4 3k Millipore) and injected into rabbits for polyclonal antibody production (see Antibodies).

For affinity purification of polyclonal anti-Klp98A antibodies, we purified His-Klp98A-Cter following a protocol similar to the one described above, but using NiNTA resin (Ni Sepharose High Performance, Amersham) and 10 mM imidazole in lysis and wash buffers. His-Klp98A-Cter was eluted by 20 mM HEPES, 150 mM KCl, 300 mM imidazole, 1 mM DTT, pH 7.7, dialysed against 20 mM HEPES, 150 mM KCl, 10% glycerol, 1 mM DTT, pH 7.7, concentrated by ultrafiltration to 7.3 mg ml<sup>-1</sup> and finally coupled to amino-link sepharose resin (Pierce).

His-GFP and His-GBP were expressed and purified from *E. coli* BL21 Rosetta 2 following the same procedure as for His-Klp98A-Cter. Final dialysis buffer was (20 mM HEPES, 150 mM NaCl, pH 7.7) for His-GFP and (20 mM HEPES, 150 mM NaCl, 5% glycerol, 15 mM imidazole, pH 7.7) for His-GBP. His-GFP and His-GBP were concentrated by ultrafiltration to 7.5 mg ml<sup>-1</sup> and 2.34 mg ml<sup>-1</sup>, respectively, flash frozen in liquid N<sub>2</sub> and kept at -80°C.

GST-tagged Patronin-Cter purification was similar to the one of GST-Klp98A, except that TEV was removed by using NiNTA resin before final dialysis. Tag-free Patronin-Cter was injected into rabbits for polyclonal antibody production. Alternatively, tag-free Patronin-Cter was coupled to amino-link sepharose resin for affinity purification these anti-Patronin antibodies (see Antibodies).

Klp98A-GFP-PC (that is, full length Klp98 fused to GFP and the PC tag in Cter) was purified from a puromycin-resistant Schneider S2 stable cell line expressing Klp98A-GFP-PC under the inducible metallothionein promoter. To obtain this cell line, S2 cells were transfected with pMT Puro Klp98A-GFP-PC plasmid (see above) using Effectene (Qiagen). This stable cell line was subsequently grown and selected in Schneider medium (Life Technologies) enriched with 10% vol/vol fetal calf serum and 5  $\mu$ g ml<sup>-1</sup> puromycin (Applchem). The concentration of inducer (CuSO<sub>4</sub>) was subsequently gradually increased from 0.05 mM to 0.6 mM over 1 month so as to select clones able to express high amounts of Klp98A (whose overexpression is toxic). We then grew 100 15-cm plates of this pseudo-clone. Cells were harvested, washed in XB buffer (20 mM HEPES, 150 mM KCl, 1 mM CaCl<sub>2</sub>, pH 7.7) then lysed in 100 ml of lysis buffer (20 mM HEPES, 150 mM KCl, 1% Triton X-100, 1 mM CaCl<sub>2</sub>, 2 mM MgCl<sub>2</sub>, 0.1 mM ATP, pH 7.2) supplemented with a protease inhibitor cocktail (benzamidine/chymostatin/antipain/leupeptine/pefabloc/PMSE, see SDS-PAGE and western blot section). Lysate was rocked for 1 h at 4°C to ensure microtubule depolymerization. Cell debris were removed by centrifugation at 3,300g for 10 min at 4°C in a swinging bucket rotor (Heraeus Megafuge) followed by an ultracentrifugation at 200,000g for 30 min at 4°C (Beckman Ti 60). Clarified lysate was subsequently incubated with 1 ml of pre-equilibrated Protein C affinity resin (Roche) for 4 h at 4°C with recirculation. The column was then washed extensively with 50 ml lysis buffer, then with 50 ml of Klp98A buffer (20 mM HEPES, 150 mM KCl, 2 mM MgCl<sub>2</sub>, 0.1 mM ATP, 10% glycerol, pH 7.2) enriched with 1 mM CaCl<sub>2</sub>, followed by 50 ml Klp98A buffer. Elution was then performed by incubating the 1 ml resin with 1 ml of Klp98A buffer enriched with 5 mM EGTA overnight at 4°C with rocking. Eluted Klp98A-GFP-PC was then mixed with Klp98A buffer enriched with 2 mM DTT in a 50:50 volume ratio, concentrated by ultrafiltration (Amicon Ultra-4 3k Millipore), and further purified by gel filtration on a Superdex 200 10/300 column (GE Healthcare Life Sciences) in (20 mM HEPES, 0.15 M KCl, 2 mM MgCl<sub>2</sub>, 1 mM DTT, 0.1 mM ATP, pH 7.2) at 0.25 ml min<sup>-1</sup>. Fractions containing Klp98A-GFP-PC were pooled, mixed with Klp98A buffer containing 20% glycerol final in a 50:50 volume ratio, concentrated by ultrafiltration (Amicon Ultra-4 3k Millipore), flash frozen in liquid N<sub>2</sub> and finally kept at -80°C (Fig. 1b). Final Klp98A-GFP-PC buffer is (20 mM HEPES, 150 mM KCl, 2 mM MgCl<sub>2</sub>, 0.1 mM ATP, 10% glycerol, 1 mM DTT pH 7.2). For motility assays were a high concentration of Klp98A-GFP-PC was critical to achieve a high density of Klp98-GFP-PC on the quantum dots, the gel filtration step was omitted.

Unlabelled porcine tubulin or HiLyte488- and rhodamine-labelled porcine tubulin were purchased from Cytoskeleton, reconstituted at 10 mg ml<sup>-1</sup> in BRB80 buffer (80 mM K-Pipes, 1 mM MgCl<sub>2</sub> pH 6.9) supplemented with 1 mM GTP (Roche) or 1 mM GMPPCP (Jena Bioscience), flash frozen in liquid N<sub>2</sub> and kept at -80°C.

GFP-MAP65-1 was a gift from V. Stoppin-Mellet, M. Vantard and J. Gaillard (ref. 13).

**Fly notum live imaging, Delta antibody uptake and SiR-tubulin imaging.** Fly notum dissection and SOP imaging was performed in clone 8 medium after embedding into a fibrinogen clot<sup>48,49</sup> in order to diminish tissue movements during fast 3D image acquisition as described<sup>50</sup>.

Fluorescent Delta antibody uptake to label the Sara endosomes was performed as previously described<sup>1,50</sup> with a 5-min pulse (3.4  $\mu$ g ml<sup>-1</sup> antibody in clone 8) and a

20-min chase (referred to as iDelta<sub>20</sub>). To address antibody bleaching, which hampers the accuracy of endosome tracking during acquisition, we replaced the original primary anti-Delta antibody coupled to a fluorescent Fab<sup>1,50</sup> by a primary anti-Delta antibody covalently coupled to the very stable Atto647N dye (see Antibodies). Under these labelling conditions, no bleaching is detectable (Extended Data Fig. 4n).

For SiR-tubulin imaging, dissected nota were incubated in clone 8 medium enriched with 1  $\mu$ M SiR-tubulin<sup>15</sup> (Spirochrome) for 30 min at room temperature, then washed twice in clone 8 before fibrinogen clot embedding as above and imaging. Note that SiR-tubulin is less excluded from the Pavarotti-positive central spindle core than Jupiter-mCherry (Fig. 3a).

For imaging of Sara endosomes dynamics *in toto* with neither iDelta uptake nor notum dissection (Extended Data Fig. 2c–h), pupae were mounted as described by Jaffred and Bellaiche<sup>51</sup>. Drift along the z axis resulting from muscle contractions was corrected by manually adjusting the focus during the acquisition. Compared to the signal in the primary culture preparation upon an antibody uptake, this *in toto* preparation shows a lower signal-to-noise ratio owing to the glow signal generated by the tissues underneath the epithelium of the epidermis. To address this, and only for visualization purposes, in Extended Data Fig. 2f, g we processed the images with a wavelet à trous filter (ImageJ plugin 'Kymo Toolbox' developed by Fabrice Cordelières).

Imaging was performed using a 3i Marianas spinning disk confocal setup based on a Zeiss Z1 stand, a 63 $\times$  PLAN APO NA 1.4 objective and a Yokogawa X1 spinning disk head followed by a 1.2 $\times$  magnification lens and an Evolve EMCCD camera (Photometrics). Fast z-stack acquisition of entire SOP cells (0.5- $\mu$ m steps) was obtained using a piezo stage (Mad City Labs). Single-emitter emission filters were always used to avoid bleed-through and each channel was acquired sequentially. To increase acquisition speed for iDelta<sub>20</sub> endosome tracking, we acquired 3D stacks spanning only 3  $\mu$ m along the z axis (with 0.5- $\mu$ m steps), which is usually sufficient to contain most of the central spindle (and sufficient to distinguish particles along the z axis, given the PSF of the microscope at this wavelength). In addition, the Pavarotti channel was acquired once every 20 time points. The strong brightness of the Atto647N dye allowed us to perform 3D acquisition at 1.3 Hz on average. Unless stated otherwise, data presented in figure panels correspond to maximum-intensity projections.

**Fly notum and S2 cells immunofluorescence.** Dissected fly nota were fixed according to a method designed to preserve the microtubule cytoskeleton<sup>52</sup>. In brief, nota were first incubated in Hank's balanced salt solution (Gibco) enriched with 1 mM DSP (Pierce) for 10 min at room temperature followed by a 10 min incubation in MTSB (microtubule stabilization buffer: 0.1 M PIPES, 1 mM EGTA, 4% PEG 8000, pH 6.9) enriched with 1 mM DSP, then finally in MTSB enriched with 4% PFA (Electron Microscopy Science). Nota were then permeabilized in MTSB enriched with 4% PFA and 0.2% Triton X-100 then processed for immunofluorescence as described<sup>1</sup> and mounted in Prolong Gold anti-fade reagent (Molecular Probes). Unlabelled and fluorescently labelled (see Antibodies) primary antibodies were used at 1  $\mu$ g ml<sup>-1</sup>. When non-labelled primary antibodies were used, we added Alexa647- and Alexa488-coupled secondary antibodies (Life Technologies) at a 1:500 dilution.

For lineage staining (Extended Data Fig. 3c), fly nota were dissected 30 h after puparium formation and processed for immunofluorescence as above using primary rat anti-Elav at 22  $\mu$ g ml<sup>-1</sup> antibodies followed by Cy5-coupled secondary antibodies (Biozol) at a 1:100 dilution.

For S2 cells immunofluorescence (Extended Data Fig. 6b), cells were plated onto glass coverslips pre-coated with Concanavalin A (Sigma, 0.05 mg ml<sup>-1</sup> in water for 1 h) for 1 h at 25°C in Schneider medium enriched with 10% serum. Cells were then fixed with 4% PFA (Electron Microscopy Science) for 20 min, then processed for immunofluorescence using standard techniques with Oregon-green 514-anti  $\alpha$ -tubulin antibodies at 1  $\mu$ g ml<sup>-1</sup> final (see Antibodies). Coverslips were mounted in Prolong Gold anti-fade reagent.

Image acquisition was performed on the 3i Spinning disk confocal microscope described above, but using a 100 $\times$  PLAN APO NA 1.45 TIRF objective and a z step of 0.27  $\mu$ m for optimal sampling along the z axis. Alternatively, for Extended Data Fig. 3c, images were taken on this setup using a 40 $\times$  PLAN APO NA 1.3 objective and a Photometrics HQ2 CCD camera.

**Fly notum preparation for co-localization analysis.** For co-localization studies of iDelta<sub>20</sub> with GFP-Sara (Extended Data Fig. 2a, b) and of Klp98-mCherry with GFP-Sara, iDelta<sub>20</sub>, GFP-Rab5 knock-in and YFP-Rab7 knock-in (Extended Data Fig. 1c–e), dissected fly nota embedded in the fibrinogen clot were fixed using 4% PFA in PEM buffer (80 mM K-Pipes, 5 mM EGTA, 1 mM MgSO<sub>4</sub>, pH 6.95) for 20 min at room temperature, then washed three times with PEM and imaged in PEM. Image acquisition was performed on the 3i Spinning disk confocal microscope described above with the 100 $\times$  PLAN APO NA 1.45 TIRF objective, a z step of 0.27  $\mu$ m and both channels were acquired sequentially at each z plane. Cells at various stages of the cell cycle were included into the analysis.



Since signal of YFP-Rab11 at endogenous levels (knock-in) was lost upon fixation in our conditions, co-localization between Klp98-mCherry and YFP-Rab11 was addressed in living tissue (acquiring only one  $z$  plane, to address fast 3D movements of the endosomes).

**Antibodies.** Polyclonal rabbit anti-Klp98A antibody was generated by injecting rabbits (Eurogentec Speedy program) with cleaved GST-Klp98A-Cter (see Protein purification). Immunized serum was subsequently affinity-purified with sepharose-bound His-Klp98A-Cter using standard glycine (0.1 M, pH 3.0) elution. Eluted antibody was subsequently dialysed against PBS then PBS-50% glycerol for storage at  $-20^{\circ}\text{C}$ . The characterization of this antibody is presented in Extended Data Fig. 1a, b. Polyclonal rabbit anti-Patronin antibody was generated using the same protocol. Its characterization is provided in Extended Data Fig. 6a.

Mouse anti-Delta monoclonal antibodies (C594.9B, Developmental Studies Hybridoma Bank) were purified on a Protein G column (Pierce) from hybridoma culture supernatant obtained by cultivating the hybridoma in CELLline devices (Integra) using RPMI medium (Gibco) supplemented with 10% ultra-low IgG fetal calf Serum (Gibco) and 1% pen-strep (Gibco). Antibodies were subsequently dialysed against fresh 0.15 M sodium bicarbonate pH 8.3, concentrated to  $4.11\text{ mg ml}^{-1}$  and labelled with NHS-Atto 647N (Atto tech) in a  $5\times$  molar excess of dye for 2 h in the dark at room temperature. Free dye was subsequently removed by gel filtration on a G-25 fine column (Sigma) in PBS. Degree of labelling was measured spectrophotometrically to be 2.6.

Oregon Green 514-labelled mouse anti- $\beta$ -tubulin (E7, Developmental studies hybridoma bank), Oregon Green 514-labelled mouse anti- $\alpha$ -tubulin (12G10, Developmental studies hybridoma bank) and Atto-647N-labelled anti- $\alpha$  K40 acetylated tubulin (C3B9, HPA Cultures) were purified and labelled in a similar fashion. Degree of labelling was measured spectrophotometrically to be 2.7 for Oregon Green 514-labelled anti- $\beta$ -tubulin, 1 for Oregon Green 514 labelled anti- $\alpha$ -tubulin, and 1.6 for Atto-647N labelled anti-acetylated tubulin.

Biotinylated GBP was obtained by *in vitro* biotinylation of purified GBP (see protein purification, or purchased from Chromotek) with EZ-Link SulfoNHS protein (Pierce) in a 1:5 ratio followed by extensive dialysis (SnakeSkin 3kD MWCO, Pierce) against PBS.

All labelled antibodies were subsequently frozen in liquid  $\text{N}_2$  and kept at  $-80^{\circ}\text{C}$ . Mouse anti-PC (clone HPC4) antibodies were from Roche. Rat anti-Elav (7E8A10) was from Developmental studies hybridoma bank. Unlabelled mouse anti- $\beta$ -tubulin (E7) was also used for loading controls in western blots.

**In vitro motility assays.** General tubulin handling as well as preparation of GMPPCP-stabilized, Taxol-stabilized and polarity marked fluorescent microtubules were performed accordingly to the protocols of the Mitchison laboratory (<http://mitchison.med.harvard.edu/protocols.html>). GTP and GMPPCP microtubules were polymerized at  $5\text{ mg ml}^{-1}$  for 20 min at  $37^{\circ}\text{C}$  in a water bath. Unpolymerized fluorescent tubulin dimers were removed by ultracentrifugation over a glycerol cushion. Motility assays of Klp98A were performed using purified full length Klp98A-GFP-PC (that is, full length Klp98A fused to GFP and the PC tag in Cter; see Protein purification).

Imaging of motility assays were performed using a 3i TIRF microscope based on a Zeiss Z1 stand equipped with a TIRF Slider 3 module. Excitation was performed with a 488 nm laser and simultaneous detection of both microtubules and quantum dots (Qdots) was performed using a Dualview device (Photometrics) equipped with a 565dcxr dichroic (Chroma) and two emission filters (520/30 and 630/50, Chroma) in front of an EMCCD camera (Cascade II 512, Photometrics) at 6.66 Hz. The motility properties of Klp98A-bound Qdots were analysed on kymographs using the ImageJ plugin 'Kymo Toolbox' developed by Fabrice Cordelières. This plugin was also used to process images from the Qdot channel with a wavelet  $\hat{a}$  trous filter for representation purposes (Supplementary Video 5). Motility of Klp98A-bound Qdots was analysed as described<sup>53</sup> with the following modifications. In brief, glass coverslips (Agar Scientific) were cleaned using a plasma cleaner (Harrick\_plasma) and assembled into a flow chamber using sticky slides (sticky-Slide VI 0.4 Luer, Ibidi). This flow chamber was connected to an Aladdin Syringe Pump (World Precision Instrument) used to change gently the solution in the chamber. The chamber was first perfused with anti-tubulin antibodies (SAP4G5, Sigma, 1/100 dilution in BRB80) for 5 min, then passivated using four chamber volumes of  $0.1\text{ mg ml}^{-1}$  PLL-PEG (Susos) in BRB80 for 5 min followed by four chamber volumes of  $0.5\text{ mg ml}^{-1}$  K-Casein (Sigma) in BRB80 for 5 min. A dilute solution of Taxol- or GMPPCP-stabilized microtubules ( $0.05\text{ mg ml}^{-1}$ , 5% labelled with HiLyte 488) were then injected and let to adhere to the antibodies for 10 min. The chamber was then washed with four chamber volumes of imaging buffer (BRB80 enriched with  $0.25\text{ mg ml}^{-1}$  K-casein, 1 mM ATP, 40 mM DTT,  $20\mu\text{g ml}^{-1}$  catalase,  $160\mu\text{g ml}^{-1}$  glucose oxydase and 40 mM D-glucose). Klp98A-GFP-PC ( $3\mu\text{M}$ ) was pre-incubated with  $1.5\mu\text{M}$  of biotinylated GBP for 5 min, before mixing in a 10:1 molar ratio with streptavidin-coated Qdots 605 (Molecular Probes). This ensured a high density of motors per Qdot, thus mimicking a bead assay, although

bead diameter is small. These Klp98A-bound Qdots were then injected in the flow chamber in imaging buffer.

Gliding assays of polarity-marked microtubules (Fig. 1c) were performed using the same flow chamber described above. Polarity-marked microtubules were obtained by elongating short bright GMPPCP microtubule seeds ( $5\text{ mg ml}^{-1}$ , 30% rhodamine labelled) with a dimmer tubulin mix ( $1.5\text{ mg ml}^{-1}$ , 5% rhodamine labelled) followed by stabilization with  $20\mu\text{M}$  Taxol. The chamber was first perfused with Klp98A-GFP-PC ( $2.9\mu\text{M}$ ) then passivated with PLL-PEG as above. Polarity-marked Taxol-stabilized microtubules were then injected and let to adhere to Klp98A for 5 min. The chamber was then washed with two chamber volumes of imaging buffer enriched with  $20\mu\text{M}$  Taxol then imaged in the same buffer. As seen in Fig. 1c, the minus-end (short) is leading in these gliding assays, indicating that Klp98A is a plus-end motor.

For motility of Klp98A-bound Qdots on antiparallel arrays of microtubules, antiparallel bundles were generated by incubating  $50\text{ nM}$  GMPPCP microtubules (5% rhodamine-labelled) with  $6.5\text{ nM}$  of GFP-MAP65-1<sup>13</sup> for 5 min at room temperature. These bundles were injected into the chamber and moving Klp98A-bound Qdots were observed as before using a 561 nm laser to excite rhodamine and a 405 nm laser to excite the Qdots 605. Due to the excess of Klp98A-GFP-PC over GBP-biotin in this assay, it is likely that all available GFP-binding sites of the Qdots are saturated, thus the presence of GFP-tagged MAP65-1 is not an issue.

For the analysis of the frequency at which Qdots change direction, we only considered antiparallel overlaps composed of two microtubules. We first identified pauses in the motility of Qdots (a pause is defined by a Qdot immobile for at least three consecutive frames, which corresponds to 0.9 s). Then we scored the incidence of changes of direction after a pause, in order to compute the frequency of direction changes.

**Liposome flotation assay.** Liposome flotation assays were performed as described<sup>45</sup> with the following modifications. Small unilamellar vesicles (SUVs) were prepared by N. Chiaruttini in BRB80 buffer by sonication in a water bath with several lipid mixtures: DOPC:DOPS 90:10; DOPC:DOPS:PI(3)P 80:10:10; DOPC:DOPS:PI(4)P 80:10:10 and DOPC:DOPS:PI(5)P 80:10:10. All lipid mixtures were doped with 0.6% rhodamine phosphatidylethanolamine (PE).  $70\mu\text{L}$  of SUVs ( $1\text{ mg ml}^{-1}$ ) were incubated with  $5\mu\text{L}$  Klp98A-GFP-PC ( $0.05\text{ mg ml}^{-1}$ ) for 30 min at room temperature. Then  $50\mu\text{L}$  of 2.5 M sucrose in BRB80 was added and gently mixed.  $100\mu\text{L}$  of this solution was poured into a polyallomer tube (Beckman Coulter), and then overlaid with  $100\mu\text{L}$  of 0.75 M Sucrose in BRB80 then with  $20\mu\text{L}$  of BRB80. This discontinuous sucrose gradient was then ultracentrifuged at 100,000 r.p.m. for 20 min in a TLA 100.4 rotor (Beckman Coulter) at  $25^{\circ}\text{C}$  with acceleration and deceleration settings set to level 5. The top  $50\mu\text{L}$  of the gradient, referred to as the 'floating fraction', was subsequently collected and liposome recovery was quantified by measuring rhodamine fluorescence using a Spectramax I3 plate reader (Molecular Devices). Equal amounts of recovered SUVs were then loaded onto a SDS-PAGE gel followed by western blot against the PC tag to analyse protein co-flotation with the SUVs. As controls, we also loaded samples devoid of liposomes as well as the input before centrifugation (Extended Data Fig. 1f).

**Scanning electron microscopy (SEM).** Flies were euthanized by exposure to diethyl ether for 20 min, then mounted on SEM holders using double-sided carbon tape (Electron Microscopy Sciences) and subsequently treated with a gold sputter coater (JFC-1200, JEOL). Imaging was performed using a JEOL JSM-6510LV scanning electron microscope operating in high-vacuum mode using a working distance of 10 mm and an acceleration of 10 kV. Alternatively, for Extended Data Fig. 3a, imaging was performed using a JEOL 7600F scanning electron microscope using a working distance of 25 mm and an acceleration of 5 kV.

**Rationale and quantification of the *Neur* and *Numb* phenotypes.** *Rationale.* Two endocytic factors play major, independent roles during asymmetric Notch signalling in the SOP: Neuralized and Numb (reviewed in ref. 8). In Neuralized mutants, cells in the lineage become neurons and, conversely, in Numb mutants they become sockets. It has previously been shown that Neuralized complete loss of function causes a full conversion of all the SOP lineage into neurons leading to a bald notum cuticle<sup>54,55</sup>. However, a partial depletion of Neuralized in the centre of the notum (*pnr > neur<sup>RNAi</sup>*) allows many sensory organs to perform asymmetric cell fate assignment and to develop, as in wild type, into structures containing at least the two external cells (shaft and socket; Fig. 1i, j, Extended Data Fig. 3a, b).

Klp98A mutants reveal that the lineages which generated bristles in *pnr > neur<sup>RNAi</sup>* need Klp98A function to perform asymmetric cell fate assignment: in Klp98A<sup>-</sup>, *pnr > neur<sup>RNAi</sup>* double mutants, these lineages failed to perform asymmetric signalling, causing the notum to be largely bald (Fig. 1i, j, Extended Data Fig. 3a, b). This was confirmed with two independent Klp98A mutants. Conversely, these two different Klp98A mutant conditions in combination with three alternative hypomorphic mutant conditions for Numb (*Numb<sup>SW</sup>/Numb<sup>2</sup>*, *Numb<sup>SW</sup>/Numb<sup>15</sup>* or *pnr-gal4 driving Numb<sup>RNAi</sup>*) all show a strong suppression (by half) of the multiple socket phenotype diagnostic of Numb mutants<sup>9</sup>



(Extended Data Fig. 3d–f). All together, these experiments demonstrate the role of Klp98A motility in Notch signalling.

**Quantification.** To quantify these cell-fate phenotypes in the SOP lineage in *Neuralized<sup>RNAi</sup>* mutants (Fig. 1i, j and Extended Data Fig. 3a, b), we manually scored in each genotype the number of organs within the region between the left and right pairs of dorso-central macrochaetes (which corresponds to the *panier* expression region) at the dissecting scope or on SEM images. To focus on lineage specification phenotypes generated by cell-fate specification failures in the SOP division, we scored lineages which generated organs composed of one-shaft/one-socket or two-shafts. In these organs, the SOP division seems to have been normal and thereby generated a pIIa (and a pIIb cell). ‘Tufts’, which are characteristic of neuralized mutant phenotype, could be caused by SOP specification defects and were therefore excluded from the analysis.

We verified that the absence of lineages generating bristles in the *pnr > neur<sup>RNAi</sup>*, *Klp98A<sup>Δ8</sup>/Klp98A<sup>Δ47</sup>* double mutant conditions are not due to an earlier, SOP specification problem. The question is whether, in the double mutant condition, the notum is bald because SOPs were specified and the lineage has all been converted into neurons or, alternatively, whether SOPs were not specified in the first place. Immunostaining with a neural specific marker (*elav*) confirmed that, below the bald cuticle, clusters of *elav*-positive neurons are present like in the control animals (Extended Data Fig. 3c).

To quantify cell-fate phenotypes in the SOP lineage in *Numb* mutants (Extended Data Fig. 3d–f), we manually scored on SEM images the number of organs showing multiple sockets (that is, Notch gain-of-function phenotype) in the dorsal-most region of the notum (between the left and right pairs of dorsocentral bristles) both in mutant and control flies and calculated the percentage of affected organs in each genotype.

**Fluorescence lifetime imaging microscopy (FLIM).** Lifetime imaging of GFP–Patronin was performed on a setup composed of an Olympus IX81 stand, a 60× NA 1.42 oil objective, a FV1000 confocal scanner head and time-correlated single-photon counting (TCSPC) hardware from Picoquant. Illumination was achieved with a pulsed 485 nm laser (Picoquant) operating at 40 MHz, and detection was performed on a gated PMA hybrid 40 detector (Picoquant) behind a 520/35 nm bandpass filter (Semrock). Data analysis was performed using SymPhotime 2.0 software (Picoquant). GFP fluorescence lifetime was fitted to a dual exponential model after deconvolution for the instrument response function (measured using fluorescein in the presence of saturating potassium iodine). The lifetime reported in images and graphs corresponds to the intensity-weighted average lifetime.

To measure the lifetime of GFP, we incubated 10 μl of TALON beads (Clontech) with 37.5 μg of purified His–GFP (see protein purification) in 10 μl clone 8 medium for 3 h at room temperature. After two washes in Clone 8 medium, we mounted the beads on a coverslip in 50 μl clone 8 and measured the intensity-weighted average lifetime in a region of interest (ROI) encompassing each bead by FLIM, followed by averaging over several beads.

Similarly, to measure the lifetime of GFP in conditions where 100% of the molecules are bound to the GFP–nanobody (GBP), we incubated 10 μl streptavidin beads (GE healthcare) with 18 μg of biotinylated GBP (see Antibodies) for 10 min at room temperature. After extensive washing of unbound GBP, the resulting GBP-bound beads were incubated with 37.5 μg of purified His–GFP in 10 μl clone 8 medium for 3 h at room temperature. After two washes in clone 8 medium, the lifetime of GBP-bound GFP was measured as above. Alternatively, we used GFP-trap beads from Chromotek, in which the GBP is directly cross-linked to beads. This gave similar values of increased GFP lifetime:  $\tau = 2.627 \pm 0.006$  ns;  $n = 15$  for the GBP-biotin/Streptavidin beads versus  $\tau = 2.678 \pm 0.004$  ns;  $n = 10$  for the GFP-trap beads (GBP-free GFP has a lifetime of  $\tau = 2.531 \pm 0.003$  ns;  $n = 29$ ). Please note that for all FLIM measurements, either of purified GFP *in vitro* or of GFP–Patronin fusion in the fly, the term GFP refers to the enhanced GFP variant (eGFP).

**Fluorescence recovery after photobleaching (FRAP).** FRAP of GFP–Patronin (Extended Data Fig. 9e) was performed on the 3i Marianas spinning disk setup described above (63× NA 1.4 oil objective) equipped with a Micropoint Photomanipulation hardware driven by Slidebook 6.0. A region of interest (ROI) was drawn onto half of the mitotic spindle, bleached, and recovery was monitored by spinning disk confocal imaging at a frame-rate of 14.3 Hz (50 ms exposure, 20 ms transfer time). Owing to the fast recovery of GFP–Patronin (timescale of few seconds), recovery was monitored in 2D (that is, one *z* plane) to maximize frame-rate.

FRAP movies were processed as follows: signal background was first removed homogeneously using a ROI outside the cell as a reference, then, bleaching was corrected homogeneously using the first frame as a reference. GFP–Patronin signal within the bleached ROI was then integrated overtime. Intensity was then normalized using the formula:

$$I_{\text{norm}}(t) = \frac{I(t) - I_{\text{postbleach}}}{I_{\text{prebleach}} - I_{\text{postbleach}}}$$

With  $I(t)$ , the integrated intensity at time point  $t$ ;  $I_{\text{postbleach}}$ , the intensity just after bleaching, and  $I_{\text{prebleach}}$  the intensity before bleaching (averaged over five time points).

Normalized intensity was then fitted to the equation:

$$I_{\text{norm}}(t) = A(1 - e^{-t/\tau})$$

In this equation,  $A$  corresponds to the immobile fraction, the half-time of recovery is provided by  $t_{0.5} = \frac{-\ln(0.5)}{\tau}$  and  $\tau$  is an estimate of the  $k_{\text{off}}$  of GFP–Patronin for mitotic spindle microtubules (assuming that diffusion is faster than binding/unbinding kinetics). Averaging the values of  $A$ ,  $t_{0.5}$ , and  $k_{\text{off}}$  for each curve gave similar results than the values obtained by fitting the average recovery:  $A = 0.90 \pm 0.01$ ,  $t_{0.5} = 1.3 \pm 0.1$  s and  $k_{\text{off}} = 0.53 \pm 0.03$  s<sup>−1</sup>,  $n = 11$  for average of the individual fits versus  $A = 0.89 \pm 0.02$ ,  $t_{0.5} = 1.31 \pm 0.03$  s and  $k_{\text{off}} = 0.53 \pm 0.01$  s<sup>−1</sup> for fit of the average curve (95% confidence intervals).

**Image analysis.** Unless otherwise specified, image analysis was performed using custom codes written for ImageJ and Matlab (Mathworks), available on request. For representation purposes, intensity was sometimes colour-coded using the Rainbow or the Red Hot lookup tables in ImageJ. Videos were edited using Adobe Premiere Pro CS6.

**Co-localization.** To automatically measure the co-localization between iDelta<sub>20</sub> and GFP–Sara (Extended Data Fig. 2), as well as the co-localization between Klp98A–mCherry and various early endosome markers (Extended Data Fig. 1c, d), we developed a custom object-based method to determine the percentage of co-localization of signals detected in two different channels. Indeed, the fact the membrane of endosomes is organized as a mosaic of domains<sup>56–58</sup> implies that the corresponding signals only partially overlap, which explain why classical co-localization methods relying on intensity correlation coefficients perform poorly in the case of endosomes. On the other hand, object-based methods rely on the segmentation of the signals in both channels followed by the measurements of the distances between all the objects: two objects are considered co-localized if the distance between their fluorescence centroid is below a certain threshold  $r_{\text{ref}}$  (ref. 59).

Current endosome segmentation methods rely on an intensity threshold for the fluorescent signal<sup>59</sup>. This is problematic when the signal intensity in different endosomes is heterogeneous (that is, to take dim endosomes into account, bright endosomes are over-segmented, and vice-versa). To avoid this issue, we adapted to 3D a threshold-free method for endosome segmentation, which is based on Gaussian fitting. In brief, signal-positive particles in both channels are first detected in 2D in each *z* plane by a 2D Gaussian fitting algorithm<sup>60</sup>, which does not rely on an intensity threshold, but rather on the fact that particles are characterized by fluorescent signals with a spatial Gaussian distribution with an offset which correspond to the local background. Then, the particles detected in each plane (2D), but corresponding to the same object in 3D, are connected based on the point spread function (PSF) of the microscope. From this, the 3D coordinates of the centroid of fluorescence of all the particle is determined in each channel.

Once this automated detection (‘segmentation’) has been performed in the two channels, the distance  $d_{AB}$  between all particles in 3D in the two channels ( $A$  and  $B$ ) are computed and compared to a reference distance  $r_{\text{ref}}$ . If  $d_{AB} < r_{\text{ref}}$ , the particles detected in the two channels do co-localize.

When considering 2D data,  $r_{\text{ref}}$  is routinely set to be the lateral resolution of the microscope  $resol_{xy}$  (ref. 59). However, in 3D, since the axial ( $resol_z$ ) and lateral ( $resol_{xy}$ ) resolutions of the microscope are not equal, the reference distance  $r_{\text{ref}}$  has to take into account the relative position of the two particles in 3D. For instance, if the two particles are on the same *z* plane, then  $r_{\text{ref}}$  has to be  $resol_{xy}$  and conversely, if the two particles are on different *z* planes, but have identical *x* and *y* coordinates, then  $r_{\text{ref}}$  has to be  $resol_z$ . Following a method implemented by Cordelières and Bolte in the ImageJ plugin JACoP 2.0 (ref. 59), we calculated  $r_{\text{ref}}$  for the 3D problem using the following equations:

$$\Phi = \arccos \frac{(x_B - x_A)}{\sqrt{(x_B - x_A)^2 + (y_B - y_A)^2}} \text{ and}$$

$$\Theta = \arccos \frac{(z_B - z_A)}{\sqrt{(x_B - x_A)^2 + (y_B - y_A)^2 + (z_B - z_A)^2}}$$

$$r_{\text{ref}} = \sqrt{(resol_{xy} \sin\Theta \cos\Phi)^2 + (resol_{xy} \sin\Theta \sin\Phi)^2 + (resol_z \cos\Theta)^2}$$

Here,  $x_A, y_A, z_A$  and  $x_B, y_B, z_B$  are the 3D coordinates of particles in channel A and B, respectively, and  $resol_{xy}$  and  $resol_z$  correspond to the lateral and axial resolutions of the microscope, respectively. For our analysis, we measured  $resol_z = 0.9 \mu\text{m}$  and  $resol_{xy} = 0.32 \mu\text{m}$  using  $0.2\text{-}\mu\text{m}$  TetraSpeck beads from Invitrogen.

Once all the particles have been detected and their co-localization state addressed (that is,  $d_{AB} < r_{ref}$ ), we measured the percentage of co-localization as the fraction of the total signal contained in particles that do co-localize, namely:

$$\text{Per cent of colocalization} = \frac{\sum \text{intensity of colocalizing particles}}{\sum \text{intensity of total particles}} \times 100$$

This measurement was then averaged between cells and compared between genotypes. Similar values of the percentage of co-localization were obtained if the fraction of co-localizing particles rather than the fraction of total intensity was considered (data not shown).

Since much of the signal of YFP-Rab11 at endogenous levels is lost upon fixation in our conditions, we measured the co-localization between Klp98-mCherry and YFP-Rab11 in living samples. We thus acquired only single planes and applied the algorithm describe above in 2D, considering  $r_{ref} = resol_{xy}$ .

**Mean square displacement analysis of endosome tracks.** All endosome tracks were recorded with a time interval of 12 s between frames. For each endosome track, a mean square displacement (MSD) analysis was performed using the MATLAB plugin MSD Analyser<sup>61</sup>. In brief, for each endosome track in data sets of different conditions, the MSD of segments of increasing duration (delay time  $t$ ) was computed ( $\text{MSD}(t) = \langle (\Delta x)^2 \rangle + \langle (\Delta y)^2 \rangle$ ) to obtain Extended Data Fig. 4a for wild type (103 tracks) and Extended Data Fig. 4b for Klp98A<sup>-</sup> (158 tracks). The 'weighted mean' of all individual MSD traces in each condition was then computed as described<sup>61</sup>:  $\text{Weighted mean MSD}(t) = \frac{\sum_{i=0}^n w_i \text{MSD}_i(t)}{\sum_{i=0}^n w_i}$ ; where  $n$  is the number

of tracks,  $\text{MSD}_i(t)$  corresponds to the MSD value of the endosome track  $i$  for the delay time  $t$ , and  $w_i$  to the number of points averaged to compute  $\text{MSD}_i(t)$  (Extended Data Fig. 4c and Extended Data Fig. 4a, b, black curve). Note that the weighted mean gives more weight to MSD curves that have greater certainty.

We fitted two fit functions to the measured weighted MSD of endosomes as a function of delay time: (i) motion with an average velocity  $v$  and a diffusive component with a diffusion  $D$  (diffusion + directed motion), which is captured by  $\text{MSD}(t) = 4Dt + v^2 t^2$ ; and (ii) simple diffusion, captured by  $\text{MSD}(t) = 4Dt$ .

While simple diffusion (that is,  $\text{MSD}(t) = 4Dt$ ) captures well the motion of Klp98A<sup>-</sup> endosomes ( $R^2 = 0.999$ ;  $D = (2.04 \pm 0.02) \times 10^{-3} \mu\text{m}^2 \text{s}^{-1}$ ; Extended Data Fig. 4c, 95% confidence interval), it poorly fits the data when considering the motion of wild-type endosomes ( $R^2 = 0.8$ ). This indicates that Klp98A is essential for the directed motility of endosomes beyond diffusion, as seen in wild type. Indeed, the 'diffusion + directed motion' fit function (that is,  $\text{MSD}(t) = 4Dt + v^2 t^2$ ) fits well the wild-type data ( $R^2 = 0.99$ ; Extended Data Fig. 4c). This fit provides an estimate for  $v = (5.75 \pm 0.12) \times 10^{-3} \mu\text{m s}^{-1}$ , while confirming the diffusion coefficient ( $D = (1.83 \pm 0.13) \times 10^{-3} \mu\text{m}^2 \text{s}^{-1}$ ; 95% confidence interval) observed in Klp98A<sup>-</sup> conditions. Furthermore, the 'diffusion + directed motion' fit function fits the Klp98A<sup>-</sup> data well ( $R^2 = 0.999$ ) only for very low values of  $v$  ( $v = (0.3 \pm 0.5) \times 10^{-3} \mu\text{m s}^{-1}$ ;  $D = (2.11 \pm 0.04) \times 10^{-3} \mu\text{m}^2 \text{s}^{-1}$ ; 95% confidence interval), confirming that most of the directed motion of wild-type endosomes is mediated by Klp98A motor function.

Since endosomes in Klp98A<sup>-</sup> mutants display simple diffusion, we used this mutant condition to independently evaluate the diffusion coefficient of endosomes by measuring the variance of the histograms of instantaneous speed  $\frac{\Delta x}{\Delta t}$  and  $\frac{\Delta y}{\Delta t}$  in both  $x$  and  $y$  dimensions. Indeed simple diffusion along the  $x$  axis is described by  $D_x = \frac{\sigma_x \Delta t}{2}$  (ref. 61), where  $\sigma_x$  is the variance of the instantaneous speed over the  $x$  axis and  $\Delta t$  is the frame-rate (here  $\Delta t = 12 \text{ s}$ ). A corresponding expression applies to the  $y$  axis. This provided an estimate of  $D_x = 0.0024 \mu\text{m}^2 \text{s}^{-1}$  (Extended Data Fig. 4d) and  $D_y = 0.0023 \mu\text{m}^2 \text{s}^{-1}$  (Extended Data Fig. 4e), confirming the results of the MSD analysis above.

**Spatio-temporal registration of movies.** In this work, we used spatio-temporal registration of movies to generate a spatio-temporal endosome density plot during SOP division (Fig. 2d, Extended Data Fig. 7a). We also used this spatio-temporal registration to obtain a density plot of different microtubule markers to study the asymmetry of the spindle (Fig. 3a, b, Extended Data Figs 5, 6 and Supplementary Video 6). In addition, time registration allowed us to average data coming from several video data sets (Figs 1f, 3e and Extended Data Figs 2i, 4x and 7f), but also to compare the timing in different figure panels (for instance Figs 1f, 3e and Extended Data Fig. 2i).

Spatial registration was performed by defining the centre of the central spindle as monitored by the Pavarotti fluorescent signal, which is also used to establish a Cartesian system of coordinates with respect to which all the other signals

(including endosome tracks and density of microtubule markers) are referred. Time registration capitalizes in the stereotypic dynamics of Pavarotti contraction which allowed us to align the timing of our data set of videos (Extended Data Fig. 4q–s). In figure panels where data sets have been registered in time, we have set registered time = 0 to the onset of anaphase B (that is, when the Pavarotti signal starts to constrict, see Extended Data Fig. 4r).

**Spatial registration (reference frame tracking).** A custom code in ImageJ (available upon request) was generated to segment the Pavarotti signal over time. This allowed us to track the Cartesian reference frame of the central spindle, defined by an origin and two axes ( $x$  and  $y$ , where the  $y$  axis is aligned with the division plane; Fig. 2a, b, Supplementary Video 3). The orientation of the  $x$  axis is defined to be anterior to posterior (pIIb to pIIa) and was determined by automatic tracking of the mRFP-Pon signal at the anterior cortex of the SOP.

In brief, the 3D stack of confocal slices in the Pavarotti channel (GFP- or mCherry-Pavarotti;  $3 \mu\text{m}$  deep,  $\Delta z = 0.5 \mu\text{m}$ ) is projected (maximum-intensity projection), then the Pavarotti-positive region is fitted by an ellipse after semi-automated thresholding. The long axis of the ellipse defines the  $y$  axis of the reference frame described above and the short axis, the  $x$  axis (see Fig. 2a). The length of the Pavarotti-positive region along each axis is determined by taking the full-width half-maximum (FWHM) of the Pavarotti signal along the two axes. For each time point, five parameters are measured: Pavarotti width (PW, size of the Pavarotti-positive region along the  $y$  axis); Pavarotti length (PL, size along the  $x$  axis);  $x_c$  and  $y_c$ , the 2D coordinates (with respect to the top/left corner of the image) of the position of the origin C of the central spindle reference frame and  $\alpha$ , the angle defined by the  $x$  axis of this reference frame and the image horizontal axis (Extended Data Fig. 4f). The anterior to posterior orientation of the  $x$  axis was determined by detecting the position of the fluorescence centroid of mRFP-Pon signal after manual thresholding.

To evaluate the accuracy of our central spindle tracking method, we applied this tracking code on movies of PFA-fixed fly notia acquired in identical imaging conditions. We calculated the deviation from the mean value of the different parameters ( $x_c$ ,  $y_c$  and  $\alpha$ ) obtained from these movies of fixed material. We considered the FWHM of the histogram of these deviations as estimates for the accuracy of the parameters (Extended Data Fig. 4g, h, i). This analysis gave an estimated accuracy for  $x_c$ ,  $y_c$  and  $\alpha$  of 49 nm, 52 nm and  $2.4^\circ$ , respectively.

**Temporal registration (PW alignment).** Since the temporal profile of the shrinking Pavarotti width (PW) is stereotypic from cell to cell, we used it to register videos in time. For each cell, we plotted the temporal dynamics  $\text{PW}_{\text{cell}}(t)$  together with that of a reference cell ( $\text{PW}_{\text{ref}}(t)$ ; Extended Data Fig. 4o). This reference cell video was arbitrarily chosen as one that spanned from anaphase to cytokinesis, the relevant phases for this work.

We then determined the time delay  $\tau_{\text{cell}}$  that needs to be applied to the cell of interest to minimize the difference, in absolute value, between the two Pavarotti temporal profiles ( $\Delta \text{Pav}$ ), that is, find the  $\tau_{\text{cell}}$  for which  $\Delta \text{Pav}(\tau_{\text{cell}}) = \sum_{t=0}^{t=t_{\text{end}}} |\text{PW}_{\text{cell}}(t + \tau_{\text{cell}}) - \text{PW}_{\text{ref}}(t)|$  is minimum (Extended Data Fig. 4o, p). We then set the initial time of each movie to be equal to  $\tau_{\text{cell}}$  thereby registering all the movies into an 'absolute time frame'. As expected, the registered PW curves collapsed ( $R^2 = 0.93$ ) if plotted all together (Extended Data Fig. 4q–s). Importantly, the registered PL curves (Pavarotti size along the  $x$  axis), which were not used in the registration process and is a parameter independent of PW contraction, also collapsed ( $R^2 = 0.8$ ; Extended Data Fig. 4t–v), validating our time registration method.

In a few cases where the Pavarotti signal was not recorded in the video (Fig. 3e, for instance), we used instead the contraction of the Jupiter signal over the  $y$  axis as a reference. Since Jupiter is excluded from the region where Pavarotti is (Fig. 3a), the absence of Jupiter ('Jupiter gap', defined as a FWHM) can be used as a proxy of the Pavarotti region. Extended Data Fig. 4w shows that the contraction of the Jupiter gap follows that of Pavarotti, thus either marker can be used to register data sets in time.

Importantly, the contraction of Pavarotti/Jupiter is unaffected in Patronin depletion, Klp10A depletion and Klp98A mutants (Extended Data Fig. 4x), thus enabling temporal registration of videos acquired in these genetic backgrounds relative to control (Fig. 3e, h and Extended Data Fig. 7f).

**Image averaging.** To generate average videos (Fig. 3a and Extended Data Fig. 5a, b and Supplementary Video 6) the Pavarotti tracking data was used to rotate and translate each image to display them in a common spatial reference frame, the centre of which is the centre of the central spindle and whose  $x$  axis is horizontal. In order to minimize rotation artefacts, rotation was performed with bicubic interpolation after image scaling by a factor of 4 (without interpolation). After time registration, frames corresponding to each time point were processed by performing homogenous background subtraction and signal normalization



(to the brightest pixel). Finally, spatio-temporally registered videos corresponding to different cells were averaged to generate the 'average video'. All these operations were performed on  $z$ -projected images generated by signal integration over the entire volume of the spindle (sum projection,  $12\mu\text{m}$  total,  $\Delta z = 0.5\mu\text{m}$ ). Images presented in Fig. 3a correspond to late cytokinesis ( $\sim 600$  s registered time, see Extended Data Fig. 4r).

Images of fixed samples (Extended Data Figs 5d–i, 6d–f) were obtained shortly before abscission, when PW and PL (Pavarotti size along  $y$  and  $x$  axes) do not change much (registered time  $> 600$  s, Extended Data Fig. 4r) and therefore our time registration method (which relies on PW dynamics) cannot be applied anymore. At this stage, we thus used tubulin or Ac-tubulin stainings that had the characteristic '8' shape pattern of late mitotic spindles. For spatial registration, we capitalized on the fact that late spindles have a well-defined elongated 8 shape, allowing image alignment by cross-correlation with a reference image, as used in structure determination from single-particle electron microscopy data<sup>62</sup>. All these operations were performed on  $z$ -projected images (sum projection,  $6\mu\text{m}$  total,  $\Delta z = 0.27\mu\text{m}$ ).

**Quantification of endosome recruitment to the central spindle.** To generate kymographs of endosome recruitment to the central spindle (Fig. 1g), we used the Pavarotti tracking data to rotate and translate each frame (as above for video averaging, but using maximum-intensity  $z$  projection in this case). Then each frame was  $y$ -projected onto its horizontal  $x$  axis and the  $y$ -projected movie was displayed as a kymograph.

To measure endosome recruitment to the central spindle (Fig. 1f, Extended Data Figs 2i and 7f), we used the Pavarotti tracking data to measure the  $i\Delta_{20}$  signal in the central spindle region over time. To quantify the  $i\Delta_{20}$  signal, images were  $z$ -projected (sum projection) after homogeneous background subtraction using a region of the cell devoid of endosomes. This  $z$  projection was then segmented using a constant manual threshold to identify the endosomes and the  $i\Delta_{20}$  intensity signal was integrated within the segmented endosomal regions. The  $i\Delta_{20}$  intensity signal was measured both in the central spindle region and the entire cell including the central spindle. The central spindle region was operationally defined on the  $x$  axis as a  $2\mu\text{m}$  region centred at the centroid of the Pavarotti region. The central-spindle-associated signal was then expressed as a percentage of the total signal present in the cell. The Pavarotti tracking data was also used for precise time registration of these movies.

**Endosome tracking and reference frame change.** Endosome tracking was performed using a custom Matlab code. In brief, the 3D stack containing the  $i\Delta_{20}$ -Atto647N signal ( $3\mu\text{m}$  deep,  $\Delta z = 0.5\mu\text{m}$ ) was  $z$  projected (maximum-intensity projection). Particles were detected using a 2D Gaussian fitting algorithm, then tracked using a modified Vogel algorithm, as previously described<sup>60</sup>. Tracks were rendered using the ImageJ plugin mTrackJ<sup>63</sup>.

To evaluate the accuracy of our endosome tracking method, we applied this tracking code on movies of PFA-fixed fly nota acquired in identical imaging conditions. As an estimate of average accuracy of their position with respect to the image frame ( $x, y$ ), we calculated the FWHM of their distribution in this fixed material (Extended Data Fig. 4j–l). This analysis showed a positional accuracy of  $57\text{ nm}$  along the  $x$  axis and  $53\text{ nm}$  along the  $y$  axis. As expected, we found that this measured positional accuracy decreases with the signal-to-noise (SNR) ratio of the particle considered (Extended Data Fig. 4m) and we thus excluded from the analysis all the particles displaying a  $\text{SNR} < 15$ . The SNR of a diffraction limited object is defined as  $\frac{I}{\sqrt{I + \sigma^2}}$ , where  $I$  is the intensity collected at the brightest pixel

of the spot and  $\sigma$  is the standard deviation of the local background<sup>64</sup>. Importantly, due to the very high photostability of our Atto-647N anti-Delta probe, the SNR ratio of endosomes, and thus their positional accuracy, does not vary significantly over time (Extended Data Fig. 4n).

Once we have determined the position of the tracked endosomes with respect to the reference frame of the image, we then expressed these coordinates into the Pavarotti Cartesian frame defined above. We did this in order to refer the motility of the endosomes with respect to the relevant structure: the Pavarotti-positive central spindle. If the endosome has the coordinates  $\begin{bmatrix} x \\ y \end{bmatrix}$  in the image reference frame, then  $\begin{bmatrix} x' \\ y' \end{bmatrix}$  corresponds to its coordinates in the central spindle reference frame. The central spindle reference frame is centred at  $\begin{bmatrix} x_c \\ y_c \end{bmatrix}$  and oriented at an angle  $\alpha$  (see above) with respect to the image reference frame (Extended Data Fig. 4f).

The coordinates in both reference frames are related by

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix} \cdot \begin{bmatrix} x - x_c \\ y - y_c \end{bmatrix} \quad (1)$$

The precision of  $x'$  and  $y'$  thus depends on the relative precision of  $x, x_c, y, y_c$  and  $\alpha$ . The variation of  $x'$  relative to  $x, x_c, y, y_c$  and  $\alpha$  is as follows

$$dx' = \frac{\partial x'}{\partial x} dx + \frac{\partial x'}{\partial y} dy + \frac{\partial x'}{\partial \alpha} d\alpha + \frac{\partial x'}{\partial y_c} dy_c + \frac{\partial x'}{\partial x_c} dx_c \quad (2)$$

In equation (1) we have  $x' = \cos(\alpha)(x - x_c) + \sin(\alpha)(y - y_c)$  so equation (2) becomes:

$$dx' = \cos(\alpha)(dx - dx_c) + \sin(\alpha)(dy - dy_c) + [\cos(\alpha)(y - y_c) - \sin(\alpha)(x - x_c)]d\alpha \quad (3)$$

Since errors are independent, an upper estimate of the accuracy of  $x'$  (worst case scenario) is thus:

$$dx' = |\cos(\alpha)|(dx + dx_c) + |\sin(\alpha)|(dy + dy_c) + |\cos(\alpha)(y - y_c) - \sin(\alpha)(x - x_c)|d\alpha \quad (4)$$

We considered an experimental data set of  $x, x_c, y, y_c$  and  $\alpha$  from a collection of 263 data points corresponding to endosome tracks close to the Pavarotti centroid, as well as the estimated accuracy by tracking endosomes and central spindles in fixed material described above ( $dx = 57\text{ nm}$ ,  $dy = 53\text{ nm}$ ,  $d\alpha = 2.4^\circ$  (0.042 rad),  $dx_c = 49\text{ nm}$  and  $dy_c = 52\text{ nm}$ ; Extended Data Fig. 4g–l). Using this data to input into equation (4), we obtained an upper bound for the average accuracy of  $dx' = 166\text{ nm}$  in the  $x$  axis, the axis relevant to the motility of endosomes on the central spindle microtubules. Note that the bidirectional movements that we observed at the central spindle (Fig. 2e and Extended Data Fig. 4y) are in the micrometre range, which is therefore one order of magnitude larger than the accuracy of our measurements.

**Endosome track analysis.** To generate spatio-temporal endosome density plots from our data set of endosome tracks (Fig. 2d and Extended Data Fig. 7a), we binned the data (time bins = 10 s and space bins =  $0.5\mu\text{m}$ ), counted the number of tracks present in each bin and displayed this information as kymograph-type of image and applied the Red Hot lookup table.

For residence time measurements (Extended Data Fig. 7d, e), subsets of 101 tracks for control and 30 for Patronin RNAi ('high-quality tracks', see also below) were selected after gap correction by manual inspection, if necessary (see Extended Data Fig. 4y for examples). Tracks were selected (i) to be long enough (200 time points on average, thereby allowing to determine residence time); (ii) to display low motility on the  $y$  axis (indicating endosomal motility on the central spindle microtubules; Fig. 2f); and (iii) to contain at least one bidirectional motility event on the central spindle (that is, side-change event). We defined a side-change event as an event where an endosome is moving from the pIIa to the pIIb side of the spindle (or vice versa), that is, when the  $x$  coordinate of the moving endosome changes sign. On average, in our selected data, we observed  $9 \pm 1$  side changes per track, which allow determination of the average residence time on each side of the central spindle.

Residence time of endosomes on each side of the spindle was measured as follows. After detection of side-change events, the time spent by endosomes in each side of the spindle between these events was computed. Owing to the  $166\text{ nm}$  precision of our tracks within the central spindle frame (see above), we excluded from this analysis the segments of the tracks between  $x = -83\text{ nm}$  and  $x = +83\text{ nm}$ , but the result did not qualitatively change if this region is considered in the analysis (data not shown).

To measure the velocity of microtubule-based-motility, we manually selected segments within our track data where the orientation of movement in the  $x$  axis was occurring prominently in one direction for at least ten time points. These segments are referred to as 'strides'. For each selected stride, we plotted  $x$  position versus time and performed a linear fit to estimate the velocity of the stride. This gave us an estimate of  $v = 0.173 \pm 0.007\mu\text{m s}^{-1}$ .

**Measurement of  $k_{\text{off}}$ ,  $k_{\text{on}}$  and transport run length of endosomes for microtubules.** To measure the off-rate ( $k_{\text{off}}$ ) of endosomes from microtubules at the central spindle, we first automatically detected, on our central spindle tracks, which segments within the tracks correspond to events of transport on microtubules ('transport segment'). We performed this track analysis on the subset of 101 control high-quality tracks (see above and Extended Data Fig. 10b for an example). The analysis is based on the study of the properties of each step (the displacement between two frames) and the correlation between successive steps.

We operationally defined a transport segment using three criteria.

(i) Instantaneous speed in each of the steps in the transport segment must be higher than  $0.15\mu\text{m s}^{-1}$ . Since the velocity of microtubule-based-motility *in vivo* is  $v = 0.173\mu\text{m s}^{-1}$ , the diffusion coefficient is low and the frame rate is high (see below), this threshold decreases considerably the probability of incorrectly identifying a step of diffusion as a transport step. Two additional criteria help decreasing further this probability.

(ii) Segments must last for at least two consecutive steps (three frames).

(iii) The orientation of the movement must be the same for all the steps in a transport segment. These two additional criteria make negligible the probability of incorrectly identifying a diffusion segment (a segment composed only of



diffusive steps) as a transport segment. Indeed, the probability that a rare fast step of diffusion is followed by yet another rare fast step in the same orientation is extremely low.

We actually estimated by performing stochastic simulations (not shown) that, with our measured value  $D = 0.0021 \pm 0.0001 \mu\text{m}^2 \text{s}^{-1}$  and for the fastest frame rate used (1.4 Hz), the probability of incorrectly identifying a diffusion-segment as a transport-segment is about  $1 \times 10^{-3}$ . We found  $k_{\text{off}} = 0.90 \pm 0.06 \text{s}^{-1}$  from exponential fits of the distribution of the duration of transport segments (95% confidence interval; see Extended Data Fig. 10c).

To estimate  $k_{\text{on}}\rho$ , we considered the track segments in between transport segments which we defined operationally as diffusion segments. We then found  $k_{\text{on}}\rho = 0.05 \pm 0.01 \text{s}^{-1}$  (95% confidence interval; Extended Data Fig. 10d). Note that, since we analyse the tracks regardless of their position within the central spindle, the value of the measured  $k_{\text{on}}\rho$  is an average of values for different microtubule densities (that is,  $k_{\text{on}}\rho = k_{\text{on}}(\rho_a + \rho_b)/2$ ).

Extended Data Fig. 10e shows the distribution of run lengths in the transport-state. To estimate the characteristic run length  $\lambda$  for the transport state, we used the method described by Thorn and Vale (ref. 65). In brief, we determined the cumulative distribution  $P(x)$  of the transport run lengths  $x$  (that is, the fraction of run lengths shorter than a given run length). We then fitted the observed cumulative distribution  $P(x)$  to the corresponding equation  $P(x) = 1 - e^{-x/x_0}$  for  $x > x_0$ , where  $x_0 = 0.4 \mu\text{m}$  is the lower limit of runs included in the fit ( $x \leq x_0$  corresponds to short runs, which are not measured with great accuracy and are thereby excluded from the analysis). The characteristic transport run length is  $\lambda = 0.31 \pm 0.01 \mu\text{m}$  ( $R^2 = 0.98$ ; 95% confidence interval).

The advantage of the Thorn and Vale procedure is that it allows us to fit the data directly without data binning. Indeed, it has been shown that performing the exponential fit directly on the binned run length distribution (like in Extended Data Fig. 10e) yields characteristic run lengths that depend strongly on the size of the bins. **Measurement of iDelta<sub>20</sub> asymmetry.** iDelta<sub>20</sub> asymmetry was measured at late stages of cytokinesis when all endosomes had departed from the central spindle, and iDelta<sub>20</sub> asymmetry had reached its maximum ( $\sim 600 \text{s}$  in registered time, see Fig. 1f and Extended Data Fig. 2i). Asymmetry was measured as follows. Endosomes were first detected by using the 2D Gaussian fitting algorithm described above. For all data sets, the same minimal fluorescence signal above local background was imposed to detect bona fide endosomes. Total intensity was then integrated for each endosome, with the local background determined by Gaussian fitting subtracted. The pIIa and the pIIb cells were then segmented manually using the Pon channel as a reference. Finally, endosomes were assigned based on their coordinates to the segmented pIIa or the pIIb regions. The total endosomal signal for each daughter cell was subsequently computed. The percentage of iDelta<sub>20</sub> in the pIIa daughter cell was then calculated as:

$$\text{Percentage of iDelta}_{20} \text{ in pIIa} = \frac{\text{iDelta}_{20} \text{ intensity pIIa}}{\text{iDelta}_{20} \text{ intensity pIIa} + \text{iDelta}_{20} \text{ intensity pIIb}} \times 100$$

We measured the percentage of iDelta<sub>20</sub> signal in the pIIa daughter cell rather than the ratio of signal between the two cells (pIIa:pIIb) since our automatic detection method sometimes did not detect any particles in one of the two daughter cell, leading to a pIIa:pIIb ratio of 0 or infinity.

Importantly, the iDelta<sub>20</sub> asymmetries measured by this method were almost identical to results obtained with our previous method based on a 3D signal integration after manual background subtraction and thresholding<sup>1,2,50</sup> (data not shown). In addition, the iDelta<sub>20</sub> asymmetry measured by this method was similar if endosome numbers or area were considered instead of endosome intensity (data not shown).

For correlative measurements of spindle asymmetry versus iDelta<sub>20</sub> endosome asymmetry, and exploration of conditions where spindle asymmetry is inverted (Fig. 4 and Extended Data Fig. 10), we rather plotted the ratio of iDelta<sub>20</sub> in pIIa, which is calculated as:

$$\text{Ratio of iDelta}_{20} \text{ in pIIa} = \frac{\text{iDelta}_{20} \text{ intensity pIIa}}{\text{iDelta}_{20} \text{ intensity pIIa} + \text{iDelta}_{20} \text{ intensity pIIb}}$$

**Measurement of spindle asymmetry in anaphase.** In this work, we measured spindle asymmetry by two methods: the 'pseudo-line-scan' method and the 'segmentation' method (illustrated in Extended Data Fig. 5c). Both methods gave similar results in live material (Extended Data Fig. 5c corresponding to the samples displayed in Fig. 3a) and in fixed samples (Extended Data Fig. 5h, i). Unless stated otherwise, the pseudo-line-scan method was used.

**Pseudo-line-scan method.** For measurements of spindle asymmetry on live material (Fig. 3b, e), we first projected  $z$  stacks containing the entire central spindle ( $6 \mu\text{m}$  depth,  $\Delta z = 0.5 \mu\text{m}$ ) using sum-intensity projection. We then segmented

the Pavarotti signal as described above (see spatio-temporal registration), which defined  $x/y$  axes of the spindle, as well as PW (Fig. 2a). Jupiter-GFP, GFP-Patronin or SiR-tubulin signal intensity was then measured along the  $x$  axis upon signal integration over the  $y$  axis within a region of interest (ROI) centred on the Pavarotti region centroid. This measurement thus conceptually resembles a line scan along the  $x$  axis of the spindle, but a rectangular ROI, rather than a line, is considered (ROI dimensions:  $10 \mu\text{m}$  on the  $x$  axis and PW on the  $y$  axis). The signal intensity over the  $x$  axis determined this way displays two peaks: one in pIIa, one in pIIb, see Fig. 3b and Extended Data Fig. 5c. This reflects the facts that these signals are excluded (at least in part) from the Pavarotti region in the middle of the central spindle (see Fig. 2a). We then measured the value of each peak and subtracted the local background (average background was determined from five pixels adjacent to the spindle). Central spindle asymmetry was computed as the enrichment of the density of the marker in the pIIb relative to the pIIa according to

$$\begin{aligned} \text{Signal enrichment in pIIb} \\ = \frac{\text{peak intensity pIIb} - \text{peak intensity pIIa}}{\text{peak intensity pIIa}} \times 100 \end{aligned}$$

Importantly, results were almost identical if a maximum intensity projection was used instead of a sum-intensity projection, and if microtubule density was measured along a line scan with a 1 pixel width instead of the entire width of the spindle by using the ROI, suggesting that spindle asymmetry is invariant along the  $y$  axis (data not shown).

For measurement of spindle asymmetry on live material (Fig. 3e), we measured this marker enrichment in pIIb at each time point and subsequently averaged these values between different videos using the time registration method described above. In cases where frame rates were not identical among videos, the spindle asymmetry values were interpolated to the correct frame rate using spline interpolation.

The kymograph of Jupiter-GFP depolymerization (Fig. 3f) was generated by plotting the pseudo-line-scan for each time point as a kymograph. We then applied the Red Hot lookup table.

For correlative measurements of spindle asymmetry versus iDelta<sub>20</sub> endosome asymmetry, and exploration of conditions where spindle asymmetry is inverted (Fig. 4 and Extended Data Fig. 10), we plotted  $\Delta$ , the normalized enrichment of microtubule density in the pIIb side, rather than the enrichment on the pIIb.  $\Delta$  is given by the formula:

$$\Delta = \frac{\text{peak intensity pIIb} - \text{peak intensity pIIa}}{\text{peak intensity pIIb} + \text{peak intensity pIIa}}$$

Note that  $\Delta$  is symmetrical when pIIb and pIIa are inverted and that  $-1 \leq \Delta \leq 1$ .

For images of fixed samples (Extended Data Figs 5d–i, 6d–f and 8f, g), we capitalized on the fact that the spindle asymmetry as a function of time remains approximately constant at late stages of cytokinesis (Fig. 3e) and therefore measurements at those stages are unlikely to be affected by incorrect time registration. We fitted the microtubule marker signal to an ellipse to obtain the  $x$  and  $y$  axes of the spindle, determined manually (in the absence of Pavarotti signal) the cytokinesis plane and measured the microtubule enrichment in pIIb as described above considering a ROI of dimensions  $10 \mu\text{m}$  on the  $x$  axis and  $0.812 \mu\text{m}$  (4 pixels) over the  $y$  axis.

**Segmentation method.** In this method, we segmented the central spindle by considering an intensity threshold above the cytosolic background and computed the average intensity in the segmented regions in the pIIa and pIIb sides (see Extended Data Fig. 5c). This second method considers the average density of the complete pool of microtubules at the central spindle. This gave comparable results to the pseudo-line-scan method (Extended Data Fig. 5c, h, i).

**Measurement of spindle asymmetry during metaphase.** To measure spindle asymmetry in metaphase (Extended Data Fig. 5n), we first projected  $z$  stacks containing the entire metaphase spindle ( $8.5 \mu\text{m}$  depth,  $\Delta z = 0.5 \mu\text{m}$ ) using maximum-intensity projection. We then drew a line between the two spindle poles, which define the mitotic plane: the plane orthogonal to this line, located in the middle distance between centrosomes. We then measured the total signal in two ROIs of  $4.6 \mu\text{m}$  (along the mitotic plane)  $\times$   $2.3 \mu\text{m}$  (along the inter-centrosome line) on each side of the mitotic plane, in the pIIa and pIIb sides. Local background was subtracted by considering an adjacent ROI in the cell outside the spindle and the two ROIs described above. The signal enrichment on the pIIb side was then computed as

$$\begin{aligned} \text{Signal enrichment in pIIb} \\ = \frac{\text{total intensity pIIb ROI} - \text{total intensity pIIa ROI}}{\text{total intensity pIIa ROI}} \times 100 \end{aligned}$$

Importantly, these ROIs do not contain the centrosomes so that spindle asymmetry measurements are not affected by centrosome asymmetry (Extended Data Fig. 5o, p).

**Measurement of centrosome asymmetry.** To measure centrosome asymmetry of different markers throughout mitosis (Extended Data Fig. 9d–h), we first projected *z* stacks containing the entire centrosome signal (6  $\mu\text{m}$  depth,  $\Delta z = 0.5 \mu\text{m}$ ) using maximum-intensity projection. We then measured the intensity of each centrosome by considering a circular ROI centred on the centrosome (1.4  $\mu\text{m}$  diameter). Local background was subtracted by considering an adjacent ROI of identical diameter. We then calculated the ratio between the pIIa and the pIIb centrosome intensities. For prophase and prometaphase, the pIIa/pIIb centrosome identity could not be assigned since spindle rotates during metaphase. We therefore measured the ratio of the brighter centrosome over the dimer.

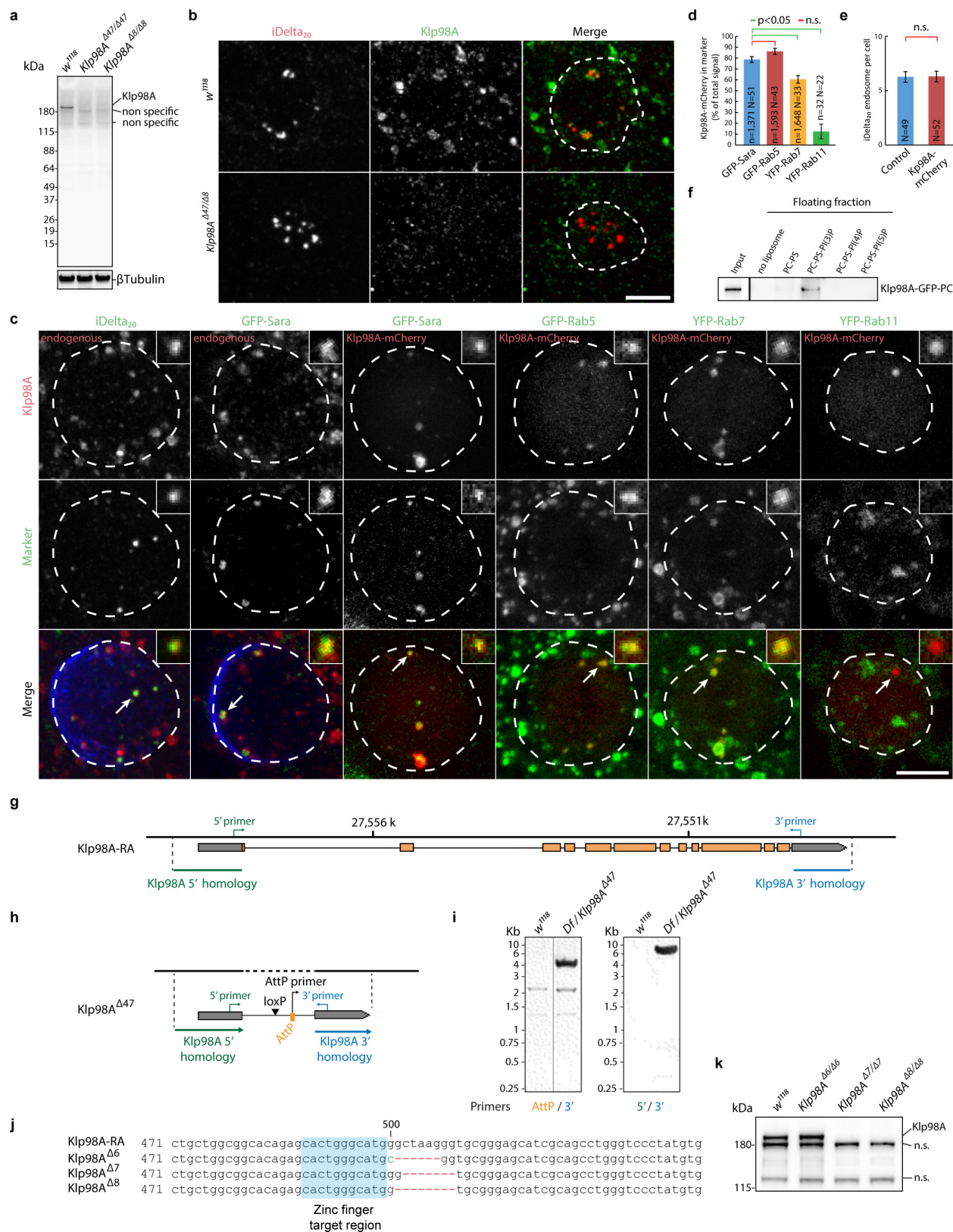
**Measurement of normalized Jupiter–GFP Density.** To compare Jupiter–GFP intensity between different videos (Fig. 3g,h), a reference intensity was needed to account for the variations of the Jupiter–GFP signal, which occurs even in identical imaging conditions and with expression of Jupiter–GFP at endogenous levels, probably owing to different imaging depths into the tissue. We decided to use the intensity of the centrosome of the pIIa daughter cell, which Jupiter labels throughout the cell cycle (Fig. 3d and Supplementary Video 7) as a reference. We measured the intensity of the pIIa centrosome by considering a circular ROI centred on the centrosome (1.2  $\mu\text{m}$  diameter) and integrating the signal intensity within the ROI on ten *z* planes (5  $\mu\text{m}$  total depth). Local background was subtracted by considering an adjacent ROI of identical diameter for each plane. We then measured the Jupiter–GFP signal in both the pIIa and the pIIb daughter cells by using the same circular ROI dimensions and background subtraction as above. We then normalized the obtained signal intensity by the pIIa centrosome value.

Interestingly, the centrosome of the pIIa daughter cell is  $1.41 \pm 0.06$  (mean  $\pm$  s.e.m.;  $n = 26$  cells) times more intense than the one of the pIIb daughter at the late cytokinesis stage considered here (Fig. 3d, Extended Data Fig. 5o, p, Supplementary Video 7). Importantly, this difference is still present in Patronin RNAi ( $1.39 \pm 0.17$ ,  $n = 24$ ) or Klp10A RNAi ( $1.25 \pm 0.08$ ,  $n = 23$ ; Extended Data Fig. 5o, p) conditions, although the values of the normalized central spindle intensities are different from wild-type conditions (Fig. 3h), suggesting that using the pIIa centrosome is indeed a good way to normalize the Jupiter–GFP data. The fact that Patronin RNAi does not affect microtubule density around the centrosome is in agreement with a recent report showing that CAMSAP2, a mammalian orthologue of Patronin, does not act on astral microtubules<sup>28</sup>.

**Statistics.** Unless stated otherwise, measurements are given in mean  $\pm$  s.e.m. Fit values (MSD analysis, Extended Data Fig. 9e and 10c–e), are provided with their 95% confidence interval. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment. No statistical methods were used to predetermine sample size. All statistical analyses were performed using SigmaStat 3.5 software (Systat) with an  $\alpha$  of 0.05. Normality of variables was verified with Kolmogorov–Smirnov tests. Homoscedasticity of variables was always verified when conducting parametric tests. For Fig. 3h, a log<sub>10</sub> transformation was applied to the data. In the case where variables failed normality and/or homoscedasticity tests, non-parametric tests were applied. In the main figures, we used Dunn's post hoc test when performing Kruskal–Wallis tests (Fig. 1h, i) and Tukey's post hoc test when performing ANOVA (Figs 3h and 4b). Post hoc tests used in Extended Data figures are indicated in their respective figure legends.

29. Gong, W. J. & Golic, K. G. Ends-out, or replacement, gene targeting in *Drosophila*. *Proc. Natl Acad. Sci. USA* **100**, 2556–2561 (2003).
30. Gong, W. J. & Golic, K. G. Genomic deletions of the *Drosophila melanogaster* *Hsp70* genes. *Genetics* **168**, 1467–1476 (2004).
31. Seum, C. *et al.* *Drosophila* SETDB1 is required for chromosome 4 silencing. *PLoS Genet.* **3**, e76 (2007).
32. Emery, G. *et al.* Asymmetric Rab 11 endosomes regulate delta recycling and specify cell fate in the *Drosophila* nervous system. *Cell* **122**, 763–773 (2005).
33. Bellaïche, Y., Gho, M., Kaltschmidt, J. A., Brand, A. H. & Schweisguth, F. Frizzled regulates localization of cell-fate determinants and mitotic spindle rotation during asymmetric cell division. *Nature Cell Biol.* **3**, 50–57 (2001).
34. Minestrini, G., Mathe, E. & Glover, D. M. Domains of the Pavarotti kinesin-like protein that direct its subcellular distribution: effects of mislocalisation on the tubulin and actin cytoskeleton during *Drosophila* oogenesis. *J. Cell Sci.* **115**, 725–736 (2002).
35. Lu, B., Ackerman, L., Jan, L. Y. & Jan, Y. N. Modes of protein movement that lead to the asymmetric localization of partner of Numb during *Drosophila* neuroblast division. *Mol. Cell* **4**, 883–891 (1999).
36. Mummery-Widmer, J. L. *et al.* Genome-wide analysis of Notch signalling in *Drosophila* by transgenic RNAi. *Nature* **458**, 987–992 (2009).
37. Bitan, A., Rosenbaum, I. & Abdu, U. Stable and dynamic microtubules coordinately determine and maintain *Drosophila* bristle shape. *Development* **139**, 1987–1996 (2012).
38. Tang, H. *et al.* Numb proteins specify asymmetric cell fates via an endocytosis- and proteasome-independent pathway. *Mol. Cell Biol.* **25**, 2899–2909 (2005).
39. Betschinger, J., Mechtler, K. & Knoblich, J. A. The Par complex directs asymmetric cell division by phosphorylating the cytoskeletal protein Lgl. *Nature* **422**, 326–330 (2003).
40. Fabrowski, P. *et al.* Tubular endocytosis drives remodelling of the apical surface during epithelial morphogenesis in *Drosophila*. *Nature Commun.* **4**, 2244 (2013).
41. Dunst, S. *et al.* Endogenously tagged rab proteins: a resource to study membrane trafficking in *Drosophila*. *Dev. Cell* **33**, 351–365 (2015).
42. Bökel, C., Schwabedissen, A., Entchev, E., Renaud, O. & Gonzalez-Gaitan, M. Sara endosomes and the maintenance of Dpp signaling levels across mitosis. *Science* **314**, 1135–1139 (2006).
43. Rothbauer, U. *et al.* A versatile nanotrapp for biochemical and functional studies with fluorescent fusion proteins. *Mol. Cell. Proteomics* **7**, 282–289 (2008).
44. Iwaki, T., Figueroa, M., Ploplis, V. A. & Castellino, F. J. Rapid selection of *Drosophila* S2 cells with the puromycin resistance gene. *Biotechniques* **35**, 482–484, 486 (2003).
45. Derivery, E. *et al.* The Arp2/3 activator WASH controls the fission of endosomes through a large multiprotein complex. *Dev. Cell* **17**, 712–723 (2009).
46. Krahn, M. P., Klopfenstein, D. R., Fischer, N. & Wodarz, A. Membrane targeting of Bazooka/PAR-3 is mediated by direct binding to phosphoinositide lipids. *Curr. Biol.* **20**, 636–642 (2010).
47. Goshima, G. & Vale, R. D. The roles of microtubule-based motor proteins in mitosis: comprehensive RNAi analysis in the *Drosophila* S2 cell line. *J. Cell Biol.* **162**, 1003–1016 (2003).
48. Forer, A. & Pickett-Heaps, J. D. Cytochalasin D and latrunculin affect chromosome behaviour during meiosis in crane-fly spermatocytes. *Chromosome Res.* **6**, 533–549 (1998).
49. Januschke, J. & Gonzalez, C. The interphase microtubule aster is a determinant of asymmetric division orientation in *Drosophila* neuroblasts. *J. Cell Biol.* **188**, 693–706 (2010).
50. Loubéry, S. & Gonzalez-Gaitan, M. Monitoring notch/delta endosomal trafficking and signaling in *Drosophila*. *Methods Enzymol.* **534**, 301–321 (2014).
51. Jauffred, B. & Bellaïche, Y. Analyzing frizzled signaling using fixed and live imaging of the asymmetric cell division of the *Drosophila* sensory organ precursor cell. *Methods Mol. Biol.* **839**, 19–25 (2012).
52. Bell, P. B. Jr & Safiejko-Mroccka, B. Improved methods for preserving macromolecular structures and visualizing them by fluorescence and scanning electron microscopy. *Scanning Microsc.* **9**, 843–857, discussion 858–860 (1995).
53. Gell, C. *et al.* Microtubule dynamics reconstituted *in vitro* and imaged by single-molecule fluorescence microscopy. *Methods Cell Biol.* **95**, 221–245 (2010).
54. Lai, E. C. & Rubin, G. M. Neuralized functions cell-autonomously to regulate a subset of notch-dependent processes during adult *Drosophila* development. *Dev. Biol.* **231**, 217–233 (2001).
55. Yeh, E., Zhou, L., Rudzik, N. & Boulianne, G. L. Neuralized functions cell autonomously to regulate *Drosophila* sense organ development. *EMBO J.* **19**, 4827–4837 (2000).
56. Gruenberg, J. The endocytic pathway: a mosaic of domains. *Nature Rev. Mol. Cell Biol.* **2**, 721–730 (2001).
57. Sönnichsen, B., De Renzis, S., Nielsen, E., Rietdorf, J. & Zerial, M. Distinct membrane domains on endosomes in the recycling pathway visualized by multicolor imaging of Rab4, Rab5, and Rab11. *J. Cell Biol.* **149**, 901–914 (2000).
58. Zerial, M. & McBride, H. Rab proteins as membrane organizers. *Nature Rev. Mol. Cell Biol.* **2**, 107–117 (2001).
59. Bolte, S. & Cordelières, F. P. A guided tour into subcellular colocalization analysis in light microscopy. *J. Microsc.* **224**, 213–232 (2006).
60. Holtzer, L. & Schmidt, T. *The tracking of individual molecules in cells and tissues* (Wiley-VCH, 2010).
61. Tarantino, N. *et al.* TNF and IL-1 exhibit distinct ubiquitin requirements for inducing NEMO-IKK supramolecular structures. *J. Cell Biol.* **204**, 231–245 (2014).
62. Frank, J. *Three-Dimensional Electron Microscopy of Macromolecular Assemblies: Visualization of Biological Molecules in Their Native State* 2nd edn (Oxford University Press, 2006).
63. Meijering, E., Dzyubachyk, O. & Smal, I. Methods for cell and particle tracking. *Methods Enzymol.* **504**, 183–200 (2012).
64. Kural, C. & Kirchhausen, T. Live-cell imaging of clathrin coats. *Methods Enzymol.* **505**, 59–80 (2012).
65. Thorn, K. S., Ubersax, J. A. & Vale, R. D. Engineering the processive run length of the kinesin motor. *J. Cell Biol.* **151**, 1093–1100 (2000).
66. Mayer, B., Emery, G., Berdnik, D., Wirtz-Peitz, F. & Knoblich, J. A. Quantitative analysis of protein dynamics during asymmetric cell division. *Curr. Biol.* **15**, 1847–1854 (2005).
67. Kirchhofer, A. *et al.* Modulation of protein properties in living cells using nanobodies. *Nature Struct. Mol. Biol.* **17**, 133–138 (2010).
68. Derivery, E., Helfer, E., Henriot, V. & Gautreau, A. Actin polymerization controls the organization of WASH domains at the surface of endosomes. *PLoS ONE* **7**, e39774 (2012).
69. Vonderheit, A. & Helenius, A. Rab7 associates with early endosomes to mediate sorting and transport of Semliki forest virus to late endosomes. *PLoS Biol.* **3**, e233 (2005).





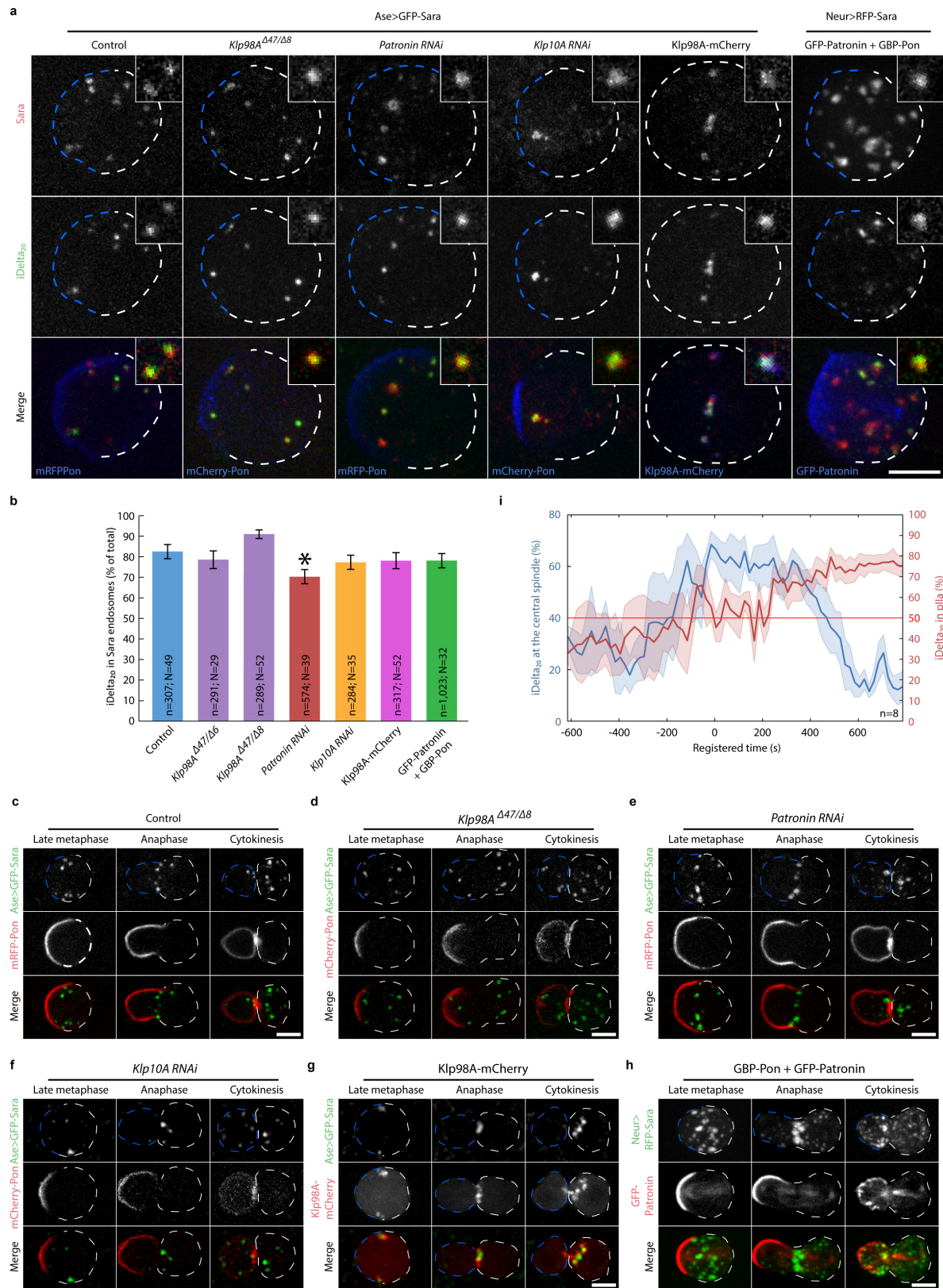
Extended Data Figure 1 | See next page for caption.



**Extended Data Figure 1 | Characterization of Klp98A. a, b,**

Characterization of the polyclonal anti-Klp98A antibody in western blot and immunofluorescence. **a**, Western blot of total extracts of *w<sup>1118</sup>* and Klp98A<sup>-</sup> mutant L2 larvae using anti-Klp98A antibodies. Klp98A is the major band above 180 kDa that disappears in Klp98A<sup>-</sup> mutants. Western blot is representative of two experiments. **b**, Confocal image showing SOPs in the notum labelled by iDelta<sub>20</sub> and Klp98A immunofluorescence in *w<sup>1118</sup>* and Klp98A<sup>-</sup> mutants. Imaging was performed by spinning-disk confocal microscopy (SDCM, z projection). Dashed line, SOP outline. The Klp98A staining associated with Sara endosomes disappears in Klp98A<sup>-</sup> mutants. **c–e**, Klp98A-containing vesicles are early endosomes. **c**, SOPs in metaphase showing Klp98A (first row; either endogenous, in the two first columns, or, in the other columns, Klp98A–mCherry at low expression levels, using temperature control) together with endosomal markers (second row; iDelta<sub>20</sub>, GFP–Sara at low expression levels, as well as GFP–Rab5 and YFP–Rab7 and YFP–Rab11 at endogenous expression levels). Third row, merge; red channel, Klp98A; green channel, endosomal marker; mRFP–Pon, in blue, is also shown in some panels. SDCM imaging. Except for YFP–Rab11, fixed samples were imaged (z projection of the entire cell; maximum intensity). Insets correspond to single planes of higher magnification images of a representative endosome (arrow). For YFP–Rab11, images correspond to one single z plane in a live cell. Dashed lines, cell outline. Note an apparent lack of fine-grain co-localization between Klp98A and iDelta<sub>20</sub>, which reflects the existence of distinct sub-domains within a particular vesicular compartment, as previously reported for early endosomes<sup>57,68,69</sup>. Scale bars in **b** and **c** are 5 μm. **d**, Automated quantification of the 3D co-localization (see co-localization section in Methods) of Klp98A with the various endosomal markers shown in **c**. Quantification is performed in 2D for YFP–Rab11. Klp98A shows a high level of co-localization with GFP–Sara and GFP–Rab5. A lower, but significant, level of co-localization is observed with YFP–Rab7, and almost no co-localization with YFP–Rab11 (Kruskal–Wallis non-parametric ANOVA followed by Dunn's post-hoc test; *N*, number of cells; *n*, number of endosomes). **e**, Quantification of the number of iDelta<sub>20</sub> endosomes per SOP in control cells and cells expressing Klp98A–mCherry at low

levels. Klp98A–mCherry expression does not affect the number of iDelta<sub>20</sub> endosomes per cell (Mann–Whitney rank sum test; *N*, number of cells). **f**, Western blot anti-Protein-C tag (PC) of the input and of floating fractions from liposome floating assays using small unilamellar vesicles (SUVs) of different compositions (DOPC:DOPS 90:10; DOPC:DOPS:PI(3)P 80:10:10; DOPC:DOPS:PI(4)P 80:10:10; DOPC:DOPS:PI(5)P 80:10:10) and purified Klp98A–GFP–PC (see Methods). Floating fractions corresponding to equal amounts of liposomes were loaded. Note that, to show the input together with the other samples in this image, the input lane is displayed after a lower exposure time and separated from the other lanes by a vertical line. Klp98A–GFP–PC directly binds to PI(3)P-containing liposomes, and does not bind to PI(4)P- nor PI(5)P-containing liposomes. Western blot is representative of two experiments. **g–i**, Generation of the *Klp98A*<sup>Δ47</sup> deletion mutant by homologous recombination (see Methods). **g**, *Klp98A* isoform A (Klp98A–RA) gene organization. Homology regions used to generate the mutant as well as PCR primers for its characterization are displayed. **h**, Gene organization after homologous recombination: the *Klp98A* gene has been replaced by an AttP site. **i**, PCR analysis of control (*w<sup>1118</sup>*) and the *Klp98A*<sup>Δ47</sup> mutant showing the expected short amplicon in the *Klp98A*<sup>Δ47</sup> mutant (5'/3' primers) and the presence of the AttP site in the *Klp98A*<sup>Δ47</sup> mutant locus (AttP/3' primers). **j**, Sequence of the three *Klp98A* alleles obtained by zinc-finger-mediated mutagenesis (see Methods). Position refers to the coding sequence of Klp98A (CG5658-PA). *Klp98A*<sup>Δ6</sup> corresponds to a missense mutation followed by a six-base-pair deletion causing missense change and deletion of two amino acids in the protein. *Klp98A*<sup>Δ7</sup> and *Klp98A*<sup>Δ8</sup> correspond to seven- and eight-base-pair deletions causing frame shifts. All the Klp98A mutants generated in this study are viable and fertile in *trans* with *Klp98A*<sup>Δ47</sup>. **k**, Western blot of total extracts of *w<sup>1118</sup>* control and Klp98A<sup>-</sup> mutant L2 larvae using anti-Klp98A antibodies. n.s., non-specific bands (see **a**) serving as loading controls. Full length Klp98A is absent in both *Klp98A*<sup>Δ7</sup> and *Klp98A*<sup>Δ8</sup> mutants, but is produced at normal levels in *Klp98A*<sup>Δ6</sup>. Western blot is representative of two experiments. For source gel images of the various western blots displayed in this figure, see Supplementary Fig. 1.



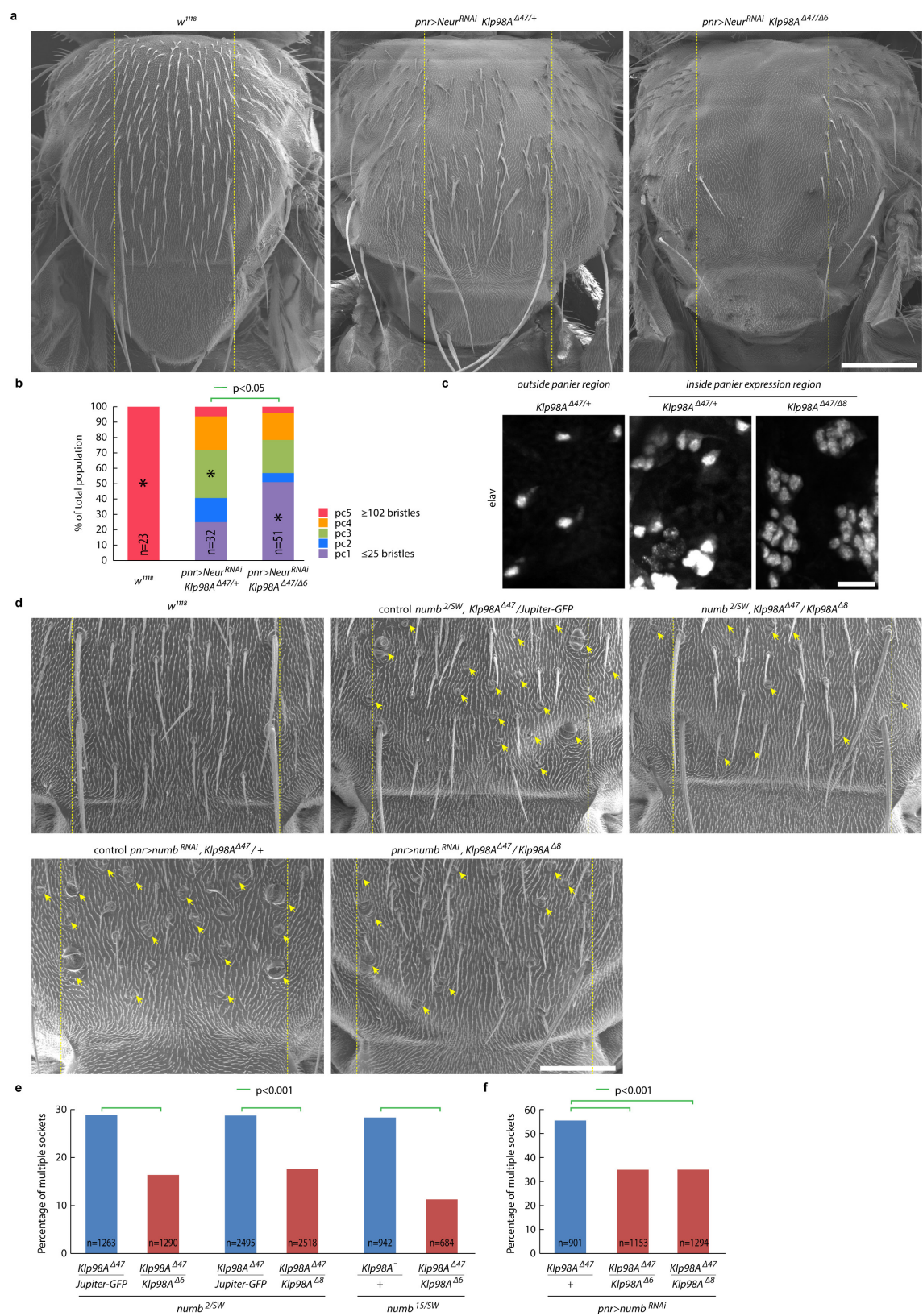
Extended Data Figure 2 | See next page for caption.

**Extended Data Figure 2 | iDelta<sub>20</sub> co-localizes with Sara and they both show the same dynamics in all the genotypes considered in this study.**

**a, b**, iDelta<sub>20</sub> co-localizes with Sara in all the genotypes considered in this study. **a**, SOPs in metaphase showing GFP–Sara or RFP–Sara (first row) and iDelta<sub>20</sub> (second row) in the different backgrounds studied in this work (SDCM imaging; image corresponds to maximum-intensity *z* projection of entire cells; insets corresponds to single planes). Third row, merge: Sara in red; iDelta<sub>20</sub> in green; mRFP–Pon, mCherry–Pon, Klp98A–mCherry or GFP–Patronin in blue. Except for the right-most panel (GBP experiment), GFP–Sara was expressed at low levels using an *ase* > GFP–Sara construct, yielding Sara endosomes which appear mostly as diffraction-limited vesicles. Dashed line, cell outline (pIIb blue, pIIa white). For the right-most panel, RFP–Sara expression is at some higher level under the control of Neur–Gal4. **b**, Automated quantification of the 3D co-localization of iDelta<sub>20</sub> and fluorescent-protein-tagged Sara in the different backgrounds (see Methods). The degree of colocalization of the iDelta<sub>20</sub> signal in Sara–FP vesicles was, as in control, around 80% amongst all the genotypes studied here (Kruskal–Wallis non parametric ANOVA followed by Dunn’s post-hoc test; *N*, number of cells; *n*, number of

endosomes). Indeed, for all the genotypes, differences in the percentage of co-localization with respect to the control were not statistically significant, with exception of *Patronin*<sup>*RNAi*</sup> (asterisk,  $P < 0.05$ ) where the co-localization is slightly lower than in control (about 12% difference). **c–h**, Dynamics of FP–Sara endosomes during SOP mitosis in all the genotypes considered in this study. Dividing SOPs showing FP–Sara endosomes in the different backgrounds studied in this report (SDCM imaging; maximum-intensity *z* projection, except for GFP–Patronin channel in **h**, which corresponds to single planes). Imaging was performed in the animal, without dissection. The dynamics of FP–Sara endosomes (spindle targeting or not, asymmetric segregation or not, and so on) parallels what we report for iDelta<sub>20</sub> endosomes throughout this study. **i**, Dynamics of the percentage of Sara endosomes at the central spindle and of endosome asymmetry measured simultaneously in control SOPs (dark line, mean; lighter area, s.e.m.). Sara endosomes become asymmetric upon spindle targeting. Time was registered between videos (see Methods) and time point zero corresponds to anaphase B onset. *n*, number of cells. Note that, for comparison, the plot of percentage of Sara endosomes at the central spindle shown here is the same as Fig. 1f. Scale bars, 5 μm.



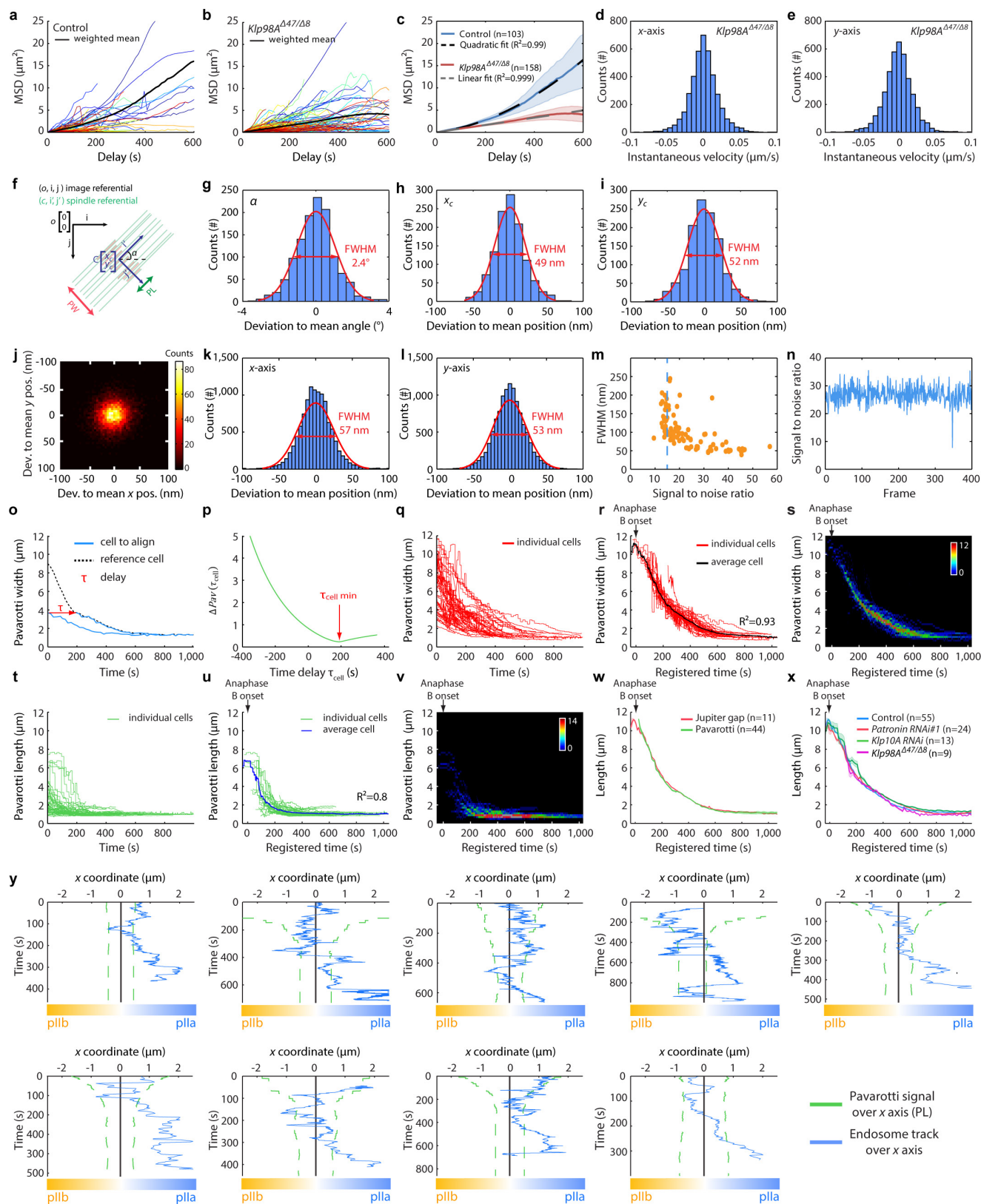


Extended Data Figure 3 | See next page for caption.

### Extended Data Figure 3 | *Klp98A* mutants affects asymmetric cell fate assignment in conditions where Neuralized or Numb are compromised.

See also 'Rationale and quantification of the *Neur* and *Numb* phenotypes' in Methods. **a–c**, *Klp98A* mutants show asymmetric cell fate assignment phenotypes in conditions where Neuralized is compromised. **a**, Scanning electron microscopy (SEM) images of adult *w<sup>1118</sup>* control, *pnr > Neur<sup>RNAi</sup> Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ6</sup>* double mutant flies and *pnr > Neur<sup>RNAi</sup> Klp98A<sup>Δ47</sup>/+* sibling controls (see 'Detailed genotypes' in Methods). Yellow dashed lines delimit the *panier* expression region corresponding to the region between the right and left pairs of dorso-central macrochaetes. For each genotype, the image corresponds to the most frequent phenotypic class in the population (asterisk in the statistical distribution shown in **b**). Scale bar, 200 μm. **b**, Number of organs in the *panier* expression region scored on SEM images for the genotypes shown in **a** (for scoring criteria, see 'Quantification of the *Neur* and *Numb* phenotypes' in Methods). Five phenotypic classes of equal range of bristles (pc1–pc5) were considered, spanning from 0 bristles (stronger phenotype) to 130 bristles (maximum number of bristles scored in control flies). n, number of flies scored. The number of lineages which generated bristles is significantly reduced in *pnr > Neur<sup>RNAi</sup> Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ6</sup>* double mutant flies compared to *Klp98A<sup>Δ47</sup>/+* heterozygous sibling controls ( $\chi^2$  test,  $P < 0.05$ ). **c**, SOPs are specified in *pnr > neur<sup>RNAi</sup> Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ8</sup>* double mutants. Neuron-specific *elav* staining of *nota* in *pnr > neur<sup>RNAi</sup> Klp98A<sup>Δ47</sup>/+* sibling controls (inside and outside the *panier* expression region) and of *pnr > neur<sup>RNAi</sup> Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ8</sup>* double mutants (inside the *panier* expression region). Note that neurons are present in *pnr > neur<sup>RNAi</sup> Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ8</sup>* double mutants, indicating that in this mutant SOPs

were specified. Scale bar, 10 μm. **d–f**, *Klp98A* mutants suppress cell fate assignment phenotypes in hypomorphic *Numb* mutants. **d**, SEM images of adult *w<sup>1118</sup>* control, *Numb<sup>2</sup>/Numb<sup>SW</sup>*; *Klp98A<sup>Δ8</sup>/Klp98A<sup>Δ47</sup>* double mutant and its *Numb<sup>2</sup>/Numb<sup>SW</sup>*; *Jupiter-GFP/Klp98A<sup>Δ47</sup>* sibling control, *pnr > numb<sup>RNAi</sup> Klp98A<sup>Δ47</sup>/+* control and *pnr > numb<sup>RNAi</sup> Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ8</sup>* double mutant. In this panel and as well as **e** and **f**, we used a *Jupiter-GFP* knock-in as a control chromosome. Arrows indicate organs with multiple sockets, characteristic of *Numb* mutants. This phenotype is suppressed in *Klp98A* mutants. Scale bar, 100 μm. Yellow dashed lines delimit the *panier* expression region where phenotypes were scored (also in the *Numb<sup>SW</sup>* conditions). **e**, Percentage of total organs showing multiple sockets in the *panier* expression region, scored on SEM images in *Numb<sup>2</sup>/Numb<sup>SW</sup>*; *Klp98A<sup>Δ6</sup>/Klp98A<sup>Δ47</sup>* double mutants (and its *Numb<sup>2</sup>/Numb<sup>SW</sup>*; *Jupiter-GFP/Klp98A<sup>Δ47</sup>* sibling controls), *Numb<sup>2</sup>/Numb<sup>SW</sup>*; *Klp98A<sup>Δ8</sup>/Klp98A<sup>Δ47</sup>* double mutants (and its *Numb<sup>2</sup>/Numb<sup>SW</sup>*; *Jupiter-GFP/Klp98A<sup>Δ47</sup>* sibling controls) and *Numb<sup>15</sup>/Numb<sup>SW</sup>*; *Klp98A<sup>Δ6</sup>/Klp98A<sup>Δ47</sup>* double mutants (and its *Numb<sup>15</sup>/Numb<sup>SW</sup>*; *Klp98A<sup>Δ47</sup>/TM6B* or *Numb<sup>15</sup>/Numb<sup>SW</sup>*; *Klp98A<sup>Δ6</sup>/TM6B* sibling controls). All double mutants show a significant reduction of the percentage of multiple sockets when compared to their respective sibling control ( $\chi^2$  test,  $P < 0.001$ ). n, total number of microchaete scored. **f**, Percentage of total organs showing multiple sockets in the *panier* expression region in *pnr > numb<sup>RNAi</sup> Klp98A<sup>Δ47</sup>/+* controls, *pnr > numb<sup>RNAi</sup> Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ6</sup>* double mutants and *pnr > numb<sup>RNAi</sup> Klp98A<sup>Δ47</sup>/Klp98A<sup>Δ8</sup>* double mutants. Both mutants show a significant reduction of the percentage of multiple sockets when compared to the control ( $\chi^2$  test,  $P < 0.001$ ).

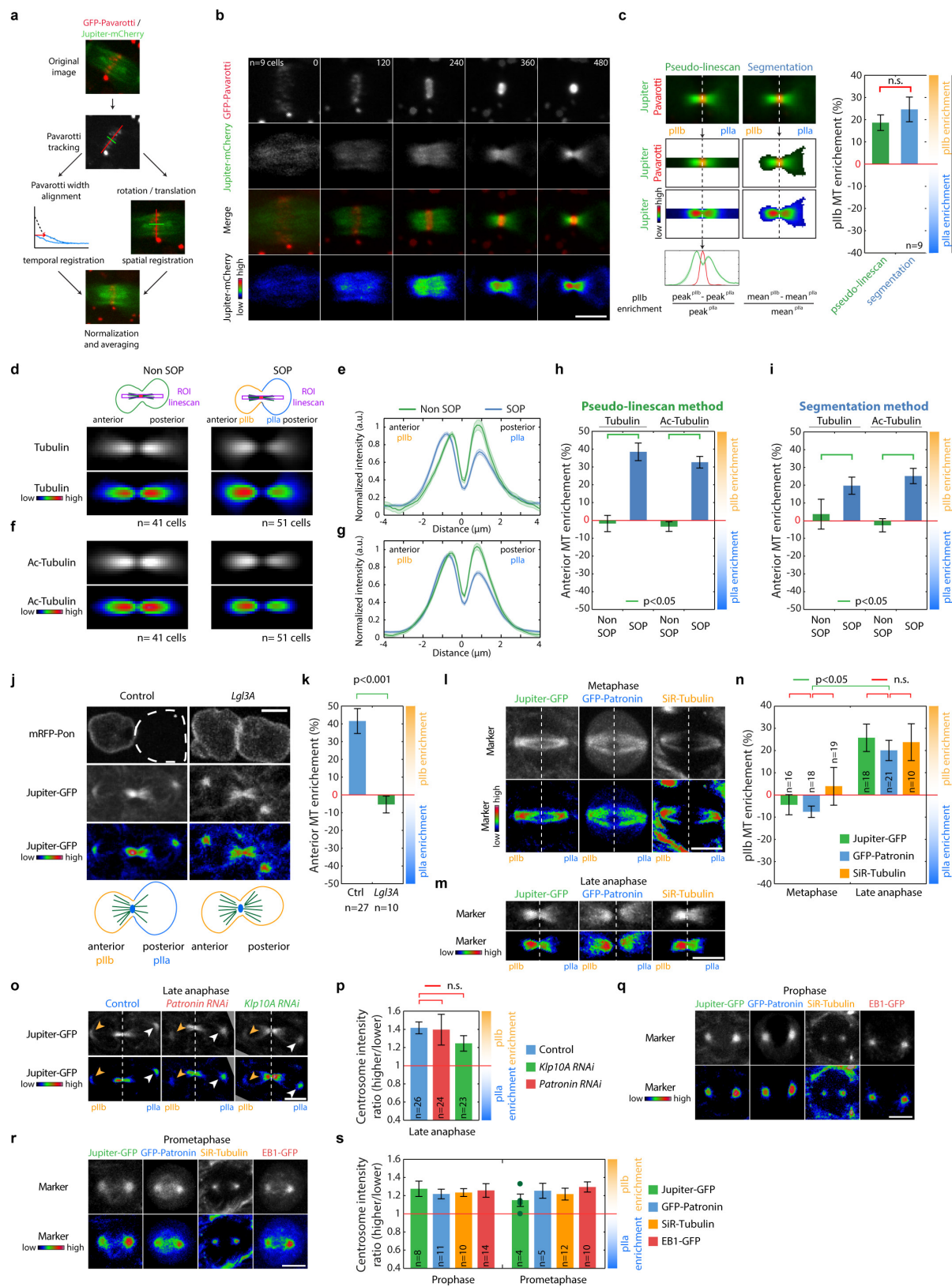


Extended Data Figure 4 | See next page for caption.



**Extended Data Figure 4 | MSD analysis of endosome tracks and precision of the tracking algorithms.** **a–e**, Determination of the endosome diffusion coefficient (see also ‘Mean square displacement analysis of endosome tracks’ in Methods). **a–b**, MSD individual traces as a function of delay time for each endosome track:  $n = 103$  tracks in control SOPs (**a**) and  $n = 158$  in *Klp98A*<sup>Δ47/Δ8</sup> SOPs (**b**). The thick black line corresponds to the weighted mean curve<sup>61</sup>, which weighs the MSD curves according to their certainty (see Methods). **c**, Parabolic or linear fitting of the weighted MSD as a function of delay time in control and *Klp98A*<sup>−</sup> SOPs (for details see Methods). Dark line, weighted mean; lighter area, s.e.m. In the absence of motor activity in the *Klp98A* null mutant, a linear fit reflects diffusive movement. In the control, a parabolic fit reflects a directed movement as a consequence of *Klp98A* motor activity. **d, e**, Histogram of instantaneous velocities of Sara endosomes in *Klp98A*<sup>Δ47/Δ8</sup> mutant SOPs along  $x$  (**d**) and  $y$  (**e**) axes. **f–i**, Precision of the Pavarotti tracking algorithm. **f**, Scheme showing the two reference frames (image reference frame and central spindle reference frame) considered in the tracking analysis. For details see ‘Spatio-temporal registration’ in Methods. **g–i**, Histograms of the deviation from the mean value of  $\alpha$  (**g**),  $x_c$  (**h**) and  $y_c$  (**i**) from videos of fixed tissue. FWHM values shown in the histograms correspond to the accuracy of each parameter. **j–n**, Accuracy of our iDelta<sub>20</sub> detection and tracking method. **j–l**, 100 iDelta<sub>20</sub> endosomes were tracked in fixed material and the deviation to their mean  $x$  and  $y$  position was computed. Deviations are represented either as a 2D density plot (**j**) or decomposed in histograms along the  $x$  axis (**k**) and the  $y$  axis (**l**). FWHM measurements indicate a positional accuracy of 57 nm on the  $x$  axis and 53 nm on the  $y$  axis, respectively. **m**, Dependence of the tracking accuracy on the endosome signal-to-noise ratio (SNR) using videos of fixed tissue (see Methods). To ensure reliable iDelta<sub>20</sub> tracking, we only considered endosomes with an SNR > 15 (dashed line). **n**, SNR of a single iDelta<sub>20</sub> endosome from a video of a fixed tissue imaged under constant laser illumination. Each time point corresponds to a 3D stack of seven slices through the central spindle. The SNR remains constant throughout iterative laser excitation of Atto647N fluorophore, indicating high

photostability of this fluorophore and, therefore, that iDelta<sub>20</sub> tracking accuracy does not decrease over time due to bleaching. **o, p**, Principle of the time registration method using the PW dynamics. Since the temporal profile of Pavarotti contraction is stereotypic, each cell can be aligned in time with a reference cell by determining the time delay  $\tau_{\text{cell}}$  that needs to be applied to this cell of interest in order to minimize the difference, in absolute value, between the two temporal profiles (**o**) (that is, find the  $\tau_{\text{cell}}$  for which  $\Delta\text{Pav}(\tau_{\text{cell}}) = \sum_{t=0}^{t=t_0} |\text{PW}_{\text{cell}}(t + \tau_{\text{cell}}) - \text{PW}_{\text{ref}}(t)|$  is minimum (**p**)). **q, t**, Unregistered PW (**q**) and PL (**t**) curves of 45 cells. **r, u**, Same data set as in **q** and **t**, but upon time registration. In this panel and others in this report, the time point zero in registered time corresponds to anaphase B onset (that is, when PW starts to contract). **s, v**, Same plot as in **r** and **u**, but represented as a density plot to show the collapse of the temporal profiles upon registration (rainbow lookup table). Note that the registered PL curves (which were not used in the registration process based on PW) also collapse, validating the method. **w**, Average temporal contraction profile of PW and ‘Jupiter gap’ (see Methods). Since Jupiter is excluded from the region that Pavarotti occupies (see Fig. 3a), the absence of Jupiter (Jupiter gap, defined as a FWHM) can be used as a proxy of the Pavarotti region. The PW and Jupiter gap temporal profiles are very similar, thus either marker can be used for temporal registration. **n**, number of cells. **x**, Average PW/Jupiter gap temporal profiles in control, *Patronin* RNAi#1, *Klp10A* RNAi and *Klp98A* mutant conditions. The profiles are very similar: time registration can be applied to *Patronin* and *Klp10A* RNAi conditions, as well as to *Klp98A* mutants. **n**, number of cells. **y**, Examples of bidirectional movements on the central spindle. iDelta<sub>20</sub> endosomes were automatically tracked during mitosis of GFP–Pavarotti expressing cells. Movement of nine endosomes on the  $x$  axis of the central spindle is shown. Green dashed line indicates PL over time. iDelta<sub>20</sub> tracks are mostly confined within the Pavarotti-positive region containing the antiparallel overlap until their departure. Tracks explore both sides of the Pavarotti positive region, but do it more often and deeper on the pIIa side.



Extended Data Figure 5 | See next page for caption.

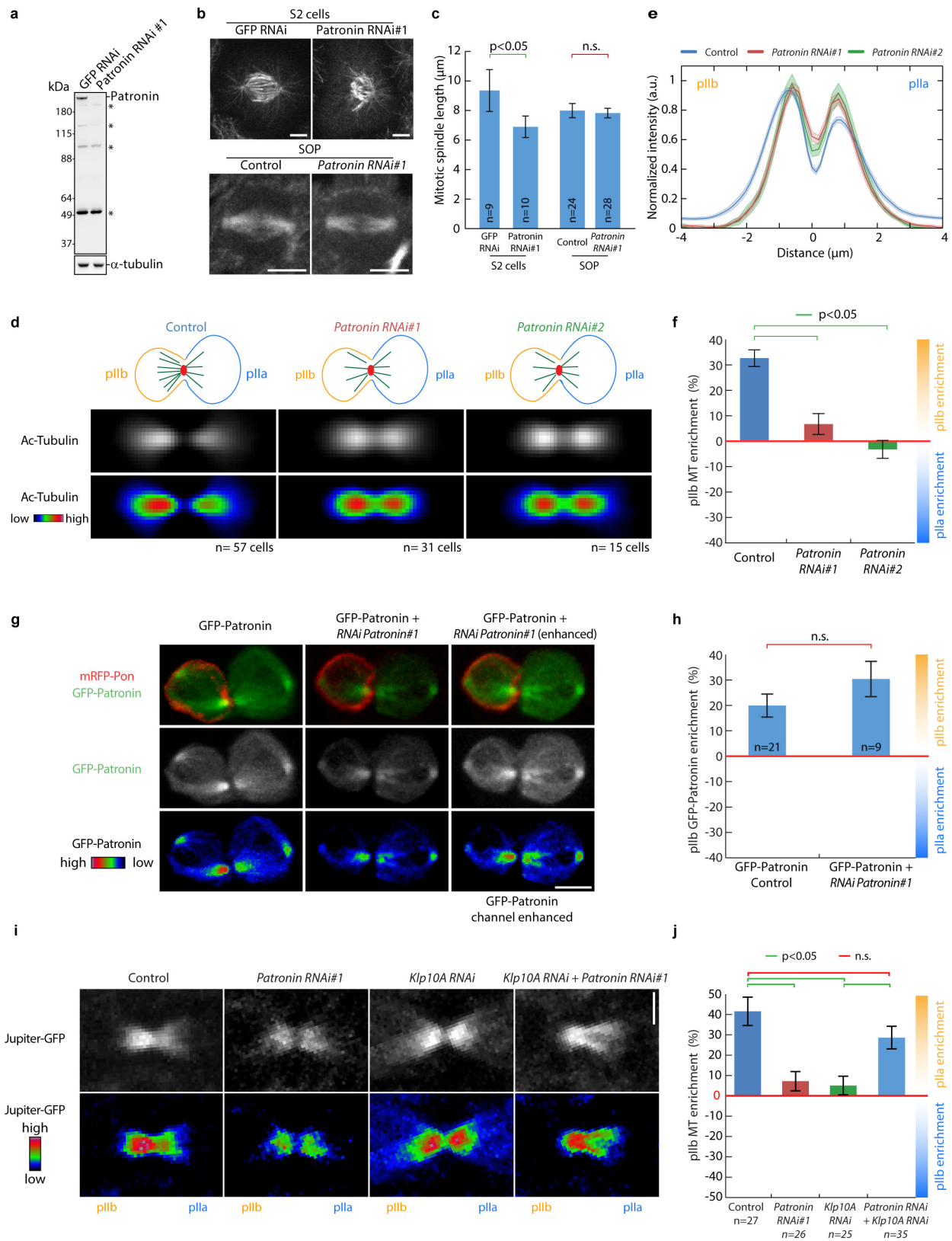
**Extended Data Figure 5 | The central spindle of SOPs is asymmetric.**

**a, b**, Generation of the 'average SOP cell'. **a**, Principle of the spindle averaging technique. Dividing SOPs expressing GFP-Pavarotti and Jupiter-mCherry were imaged by SDCM (z projection). Pavarotti signal was subsequently tracked to register all the movies in space and time before averaging. **b**, Time frames from a time-lapse average video of nine dividing SOPs. The central spindle is asymmetrical, with an enrichment of Jupiter density on the pIIb side (left). Elapsed time in seconds. **c–i**, The central spindle of SOPs is asymmetric. **c**, Left two panels: illustration of the two methods (pseudo-line-scan and segmentation method) used to measure central spindle asymmetry (see 'Measurement of spindle asymmetry in anaphase' in Methods). Right panel, Jupiter-mCherry enrichment in pIIb as scored with either method. Samples are the same as in Fig. 3a. Both methods gave comparable results when applied to the samples in Fig. 3a, but also to immunofluorescence of endogenous tubulin in fixed samples (see **h, i**). **d–g**, Average central spindle images (**d, f**) and line scan (**e, g**; line scan is based on the ROI shown in the top panels in **d** of both SOP and neighbouring epithelial cells (non-SOP) immunolabelled using Oregon Green-514 anti- $\beta$ -tubulin antibodies and Atto647N anti-acetylated- $\alpha$ -tubulin antibodies. SDCM imaging with sum projection to generate an average cell (see 'Image averaging' chapter in Methods). Anterior to the left as determined by the mRFP-Pon signal (not shown). Central spindles of SOPs are asymmetric, with an enrichment of tubulin and a corresponding enrichment of acetylated-tubulin on the pIIb side of the spindle, while non-SOP central spindles are symmetric. **h, i**, Tubulin and acetylated tubulin enrichment in the anterior pIIb cell in SOPs and anterior cells in non-SOP mitoses in the samples presented in **d** and **f** using either the pseudo-line-scan or the segmentation methods as in **c**. Note that the non-SOP cells do not show a statistically significant enrichments (zero enrichment is within the s.e.m. of the samples), while SOP cells do. Confirming this, a Kruskal–Wallis non-parametric ANOVA followed by Dunn's post hoc test ( $P < 0.05$ ) shows that in SOP cells the anterior (pIIb) enrichment of microtubule markers is higher than in non-SOP cells. These statements are true whether the line scan or the segmentation methods are considered.  $n = 47$  for non-SOP and 57 for SOP cells. **j, k**, Spindle asymmetry occurs downstream of Par complex signalling. **j**, Representative dividing control or *lgl3A*-expressing SOPs showing mRFP-Pon and Jupiter-GFP. Upon *lgl3A* overexpression, Pon segregation becomes symmetrical. Central spindle asymmetry (Jupiter-GFP channel) is also abolished upon *lgl3A* expression. Note that the control panel is the same as Fig. 3g, shown here for comparison. **k**, Jupiter-GFP enrichment in anterior pIIb daughters (control) or in anterior daughters (*lgl3A*) at late cytokinesis.  $P < 0.001$  (Mann–Whitney test). **l–s**, Analysis of spindle asymmetries from early- to late-mitotic phases (see also 'Centrosome asymmetry' in the Supplementary Discussion).

**l–n**, Spindle asymmetry occurs late during SOP mitosis. Dividing SOPs showing Jupiter-GFP, GFP-Patronin or SiR-tubulin in metaphase (**l**) or late anaphase (**m**, ~600 s registered time). Antero-posterior orientation is based on the mRFP-Pon cortical signal (not shown). **n**, Jupiter-GFP, GFP-Patronin or SiR-tubulin enrichment in the pIIb side of the spindle ( $100 \times \frac{\rho_b - \rho_a}{\rho_a}$ ;  $\rho_a$ , pIIa density;  $\rho_b$ , pIIb) in metaphase or late anaphase.

While the spindle is not asymmetric in metaphase, in late anaphase, the density of Jupiter-GFP, GFP-Patronin and SiR-tubulin is enriched in the pIIb side compared to pIIa. Statistics were performed using a Kruskal–Wallis non-parametric ANOVA followed by Dunn's post hoc test ( $P < 0.05$ ). **o, p**, Centrosome asymmetry in anaphase is unaffected upon Patronin or Klp10A depletion. **o**, SOPs in late anaphase showing Jupiter-GFP in control, Patronin-depleted or Klp10A-depleted conditions. The centrosome in pIIa (white arrow) is brighter than that in pIIb (orange arrow) in all conditions. **p**, Intensity ratio of Jupiter-GFP signal in the pIIa centrosome over the pIIb centrosome in control, Patronin- and Klp10A-depleted SOPs in late anaphase (mean  $\pm$  s.e.m.). In control, the enrichment of Jupiter-GFP in the pIIa centrosome is statistically significant (the s.e.m. does not cross the ratio = 1 line). In Patronin and Klp10A depletion this asymmetry is not statistically different to that of control (Kruskal–Wallis non-parametric ANOVA followed by Dunn's post hoc test). Note that in control cells, in the central spindle, it is the anterior pIIb (not the posterior pIIa) side that shows a higher density of microtubules. This observation untangles the asymmetry of the centrosomes from the asymmetry of the central spindle. Consistent with this, in Patronin depletion and Klp10A depletion conditions, central spindle asymmetry is abolished (see here, but also Fig. 3e, g, h), while centrosome asymmetry is not affected. **q–s**, Centrosome asymmetry can be detected before anaphase. Dividing SOPs showing Jupiter-GFP, GFP-Patronin, SiR-tubulin and EB1-GFP in prophase (**q**) or prometaphase (**r**). All four markers show a signal enrichment in one centrosome compared to the other, suggesting different microtubule-organizing activities of the two centrosomes. **s**, Intensity ratio of the centrosome-associated signals of Jupiter-GFP, GFP-Patronin, SiR-tubulin and EB1-GFP in prophase and prometaphase. Here, the pIIa/pIIb identity of the centrosome could not be determined since the mitotic spindle rotates in metaphase. We thus plotted the ratio of the brightest over the dimer centrosome. This suggests that centrosome asymmetry is acquired early, perhaps due to centrosome age. However, it must be noted that centrosome asymmetry is untangled from central spindle asymmetry. Individual data points are figured when number of cell scored is inferior to 5. All images in this figure correspond to SDCM imaging (maximum-intensity projection). Scale bars, 5  $\mu$ m. Lower panels in **b–d, f, j–m, o, q, r** correspond to rainbow lookup table.  $n$ , number of cells scored.



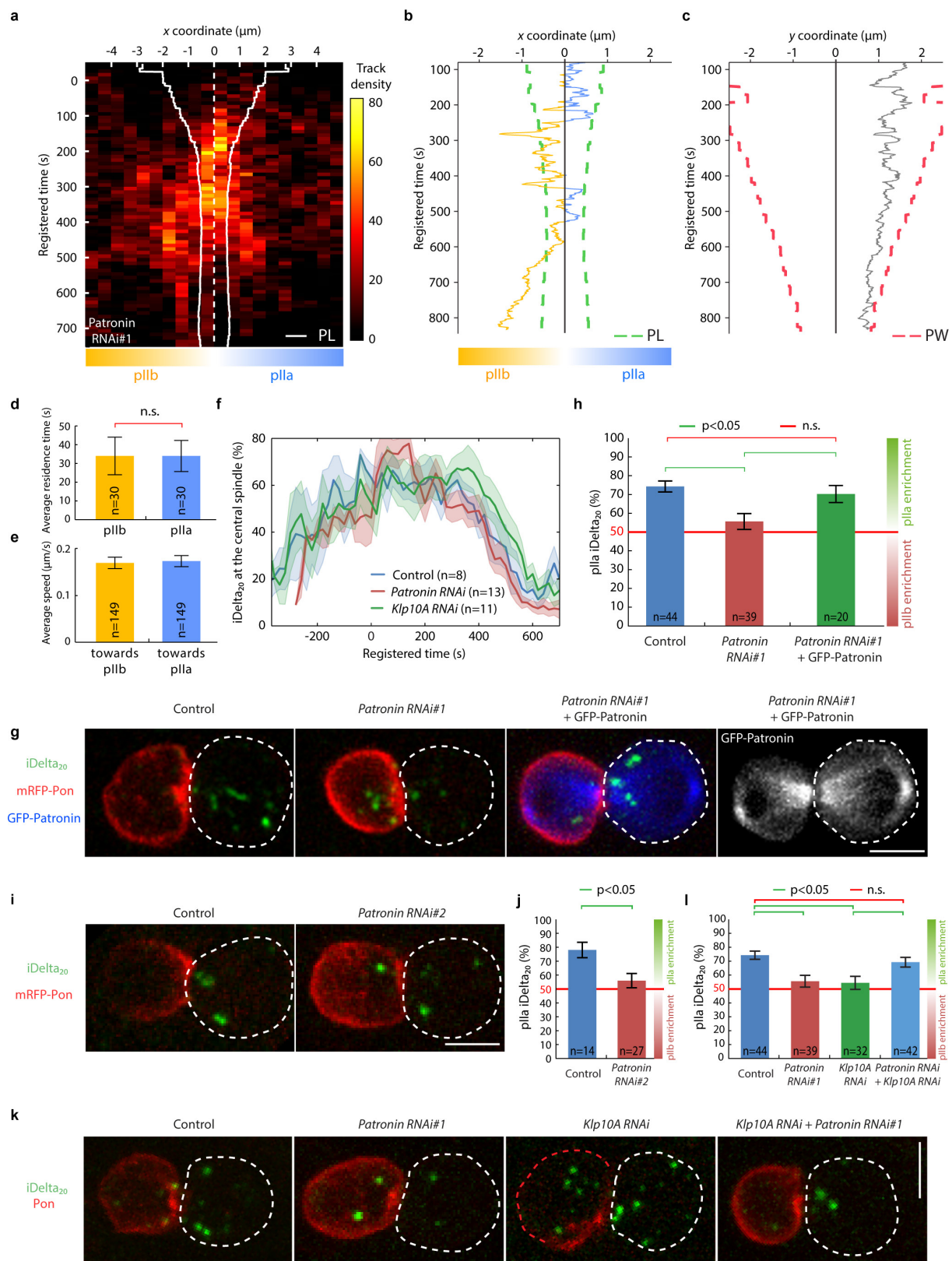


Extended Data Figure 6 | See next page for caption.

**Extended Data Figure 6 | Patronin controls central spindle asymmetry.**

**a**, Validation of the anti-Patronin antibody and of the depletion of Patronin by the *Patronin* RNAi#1 sequence. Anti-Patronin western blot of total extracts of S2 cells treated with *GFP* RNAi or *Patronin* RNAi#1. Patronin corresponds to the major band above 180 kDa that disappears upon RNAi treatment, validating both the RNAi sequence and the antibody. The *Patronin* RNAi construct presented in the main text corresponds to this *Patronin* RNAi #1. Asterisks indicate non-specific bands. Western blot is representative of two experiments. For source gel images, see Supplementary Fig. 1. **b, c**, Validation, in S2 cells, of the *Patronin* RNAi#1 sequence used in the SOP. **b**, Upper panel,  $\alpha$ -tubulin immunofluorescence of S2 cells treated with *GFP* RNAi or *Patronin* RNAi#1 in metaphase. Lower panel, Jupiter-GFP signal control and *Patronin* RNAi#1 depleted SOPs in metaphase. Live imaging. **c**, Measurement of the spindle pole to spindle pole distance of metaphase spindles in the samples described in **b**. Statistics were performed using a Kruskal-Wallis non-parametric ANOVA followed by Dunn's post hoc test;  $P < 0.05$ . Similar to what was reported by Goodwin and Vale (ref. 16) using a different RNAi sequence, depletion of Patronin in S2 cells by our *Patronin* RNAi#1 leads to short and disorganized mitotic spindles. This is however not the case in the SOP: the mitotic spindle is neither short nor disorganized. The difference between the SOP and S2 cells can be explained by the different levels of expression of the RNAi in the two systems: in the SOP, a milder Patronin depletion does allow the cell to proceed to anaphase without metaphase spindle defects/arrest. Later, the SOP central spindle seems to be more sensitive than the metaphase spindle and it is affected by our conditions of Patronin depletion in the SOP: it becomes symmetric (see below and Fig. 3e, g, h). **d, e**, Average central spindle images (**d**) and line scan (**e**; line scan is based on a ROI, as shown in Extended Data Fig. 5c) of both control and Patronin-depleted SOPs immunolabelled using Atto647N anti-acetylated- $\alpha$ -tubulin antibodies. SDCM imaging with sum projection to generate an average cell (see 'Image averaging' in Methods). Anterior to the left as determined by the mRFP-Pon signal (not shown). Spindle asymmetry is abolished upon Patronin depletion by two independent RNAi sequences. *Patronin* RNAi#1 is referred as *Patronin* RNAi in the main text. Note that the control line scan is the same as Extended Data Fig. 5g, shown here for comparison. **f**, Acetylated tubulin enrichment in the pIIb cell in control and Patronin-depleted SOPs measured by the pseudo-line-scan method (see Extended Data Fig. 5c). pIIb microtubule enrichment is defined as  $100 \times \frac{(I_b - I_a)}{I_a}$ ,

with  $I_b$  and  $I_a$  the respective pIIb and pIIa peak values of the linescans shown in **e**. The asymmetry of the central spindle is significantly reduced upon Patronin depletion (Kruskal-Wallis non-parametric ANOVA followed by Dunn's post hoc test;  $P < 0.05$ ). Note that the control data set is the same as Extended Data Fig. 5h, shown here for comparison. **g, h**, Specificity of the central spindle symmetry phenotype by *Patronin* RNAi#1 as shown in Fig. 3e, g, h: rescue of the effects of RNAi-induced Patronin depletion by overexpression of GFP-Patronin (**g**, Representative dividing SOPs expressing mRFP-Pon and GFP-Patronin (GFP-Patronin control; left panel) or, in addition, expressing *Patronin* RNAi#1 (GFP-*Patronin* + *Patronin* RNAi#1; middle and right panels) at late cytokinesis. SDCM imaging,  $z$  projection. RNAi-mediated depletion is rescued by GFP-Patronin overexpression mediated by the UAS/Gal4 system. Note that, in the RNAi conditions, the GFP-Patronin signal is reduced, reflecting the fact that the GFP-Patronin mRNA is targeted by RNAi; see, for example, GFP-Patronin signal in central versus left panels. Spindle asymmetry is rescued upon overexpression of GFP-Patronin in the RNAi condition (see right panel, where the brightness has been enhanced to visualize the different microtubule levels of the spindle). **h**, GFP-Patronin enrichment in the pIIb cell in control and rescued SOPs as in **g**. Enrichment in both conditions is not statistically significant ( $P = 0.175$ ; Mann-Whitney test). **i, j**, Co-depletion of Klp10A and Patronin suppresses their respective phenotype on central spindle asymmetry. **i**, Central spindle in dividing control, Patronin-, Klp10A-depleted and Patronin/Klp10A-co-depleted SOPs expressing Jupiter-GFP (SDCM,  $z$  projection). Antero-posterior orientation is based on the mRFP-Pon signal (not shown). Note that, upon co-depletion of Patronin and Klp10A, the phenotype of symmetric central spindle observed in *Patronin* and *Klp10A* RNAi conditions is suppressed to recover the levels of asymmetry observed in the control situation (see lookup table image; bottom panels). **j**, Quantification of GFP-Jupiter signal enrichment in the pIIb cell in control and depleted/co-depleted SOPs as in **i**. The asymmetry of the central spindle is significantly reduced upon Patronin or Klp10A depletion, but restored to control levels upon co-depletion of Patronin and Klp10A (Kruskal-Wallis non-parametric ANOVA followed by Dunn's post hoc test;  $P < 0.05$ ). Note that the control data set is the same as Extended Data Fig. 5k, shown here for comparison. Lower panels in **d, g** and **i** correspond to the rainbow lookup table. n, number of cells scored. Scale bars, 5  $\mu$ m (**b, g**) and 2  $\mu$ m (**i**).

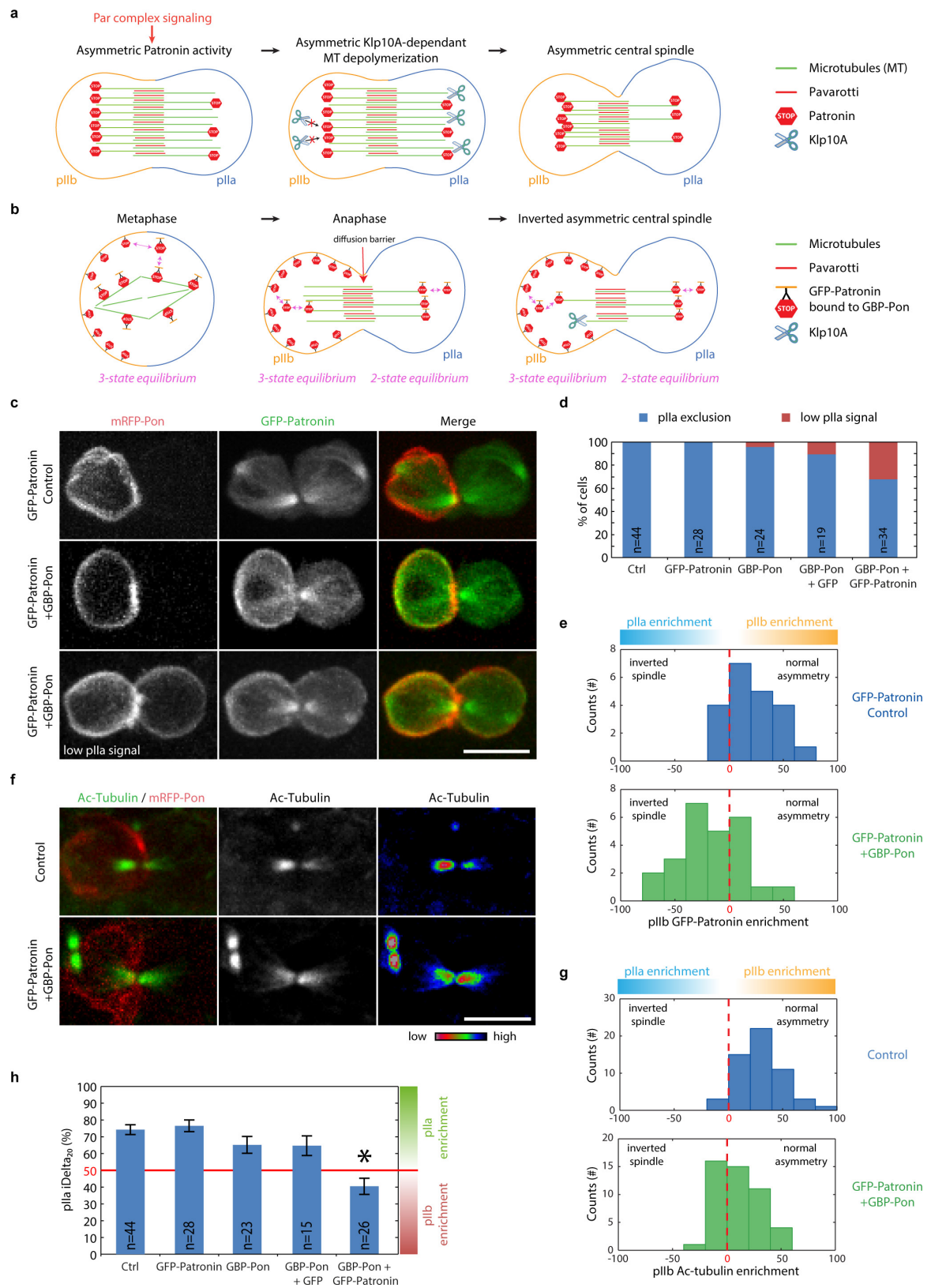


Extended Data Figure 7 | See next page for caption.



**Extended Data Figure 7 | Impaired spindle asymmetry and asymmetric Sara endosome segregation upon Patronin depletion.** **a–e**, Endosome motility analysis in *Patronin* RNAi#1 condition. **a**, Spatio-temporal density plot of iDelta<sub>20</sub> endosomes at the central spindle in Patronin-depleted cells. iDelta<sub>20</sub> endosomes were automatically tracked during mitosis of ten cells expressing GFP-Pavarotti and *Patronin* RNAi#1. After spatial registration of each track with respect to its own central spindle reference frame and temporal registration using PW constriction, the resulting 1,541 tracks were all plotted as a spatio-temporal density map (*x* axis, time). Overlaid continuous white line, averaged Pavarotti length (PL); dashed line, centroid of the Pavarotti region. iDelta<sub>20</sub> tracks are enriched within the Pavarotti-positive region. In contrast to control, tracks explore approximately equally both sides beyond the Pavarotti-positive region. **b, c**, Representative endosome track, decomposed on the *x* (**b**) and *y* (**c**) axes. On the *x* axis, movement is bidirectional and mostly confined within the Pavarotti-positive region. On the *y* axis, movements are limited or at most follow the contraction of the PW. **d**, Velocity of microtubule-based motility towards the pIIa or the pIIb cells during bidirectional movement on the *x* axis. Speed does not depend on orientation of the movement (*n* = 149 strides in 45 tracks; non-significant in Mann–Whitney test). **e**, Residence time of endosomes within both cells during bidirectional movement. In contrast to the control, there is no bias (*n* = 30 endosome tracks; non-significant in Mann–Whitney test). **f**, Dynamics of the percentage of iDelta<sub>20</sub> endosomes at the central spindle in control, Patronin- and Klp10A-depleted SOPs as in Fig. 1f (dark line, mean; lighter area, s.e.m.). iDelta<sub>20</sub> endosomes reach the central spindle in Patronin- and Klp10A-depleted cells with similar kinetics than in control cells. Note that the control data set is the same as Fig. 1f, shown here for comparison. **g, h**, Specificity of the endosomal targeting symmetry phenotype induced by *Patronin* RNAi#1 as shown in Fig. 4a, b: overexpression of GFP-Patronin rescues the effects of *Patronin* RNAi#1 on Sara endosome asymmetric segregation. **g**, Representative dividing SOP showing

mRFP-Pon and iDelta<sub>20</sub> (control), or, in addition, expressing *Patronin* RNAi#1, or *Patronin* RNAi#1 and GFP-Patronin (GFP-Patronin + *Patronin* RNAi#1, third and fourth panels) at late cytokinesis. SDCM imaging, *z* projection. Sara endosome symmetric segregation induced by *Patronin* RNAi#1 expression is rescued by overexpression of GFP-Patronin. Note that spindle asymmetry is also rescued (fourth panel), as shown above in Extended Data Fig. 6g, h. **h**, Percentage of iDelta<sub>20</sub> in the pIIa daughter cell at late cytokinesis in control, *Patronin* RNAi#1-treated and rescued cells (GFP-Patronin + *Patronin* RNAi#1). GFP-Patronin expression rescued the symmetric iDelta<sub>20</sub> segregation induced by the *Patronin* RNAi#1 (ANOVA test followed by Holm–Sidak post hoc test). **i, j**, Confirmation of the effects of Patronin depletion on Sara endosome asymmetric segregation by an independent RNAi sequence, *Patronin* RNAi#2. **i**, Representative control or *Patronin* RNAi#2-treated SOPs showing mRFP-Pon and iDelta<sub>20</sub> at late cytokinesis. SDCM imaging, *z* projection. **j**, Percentage of iDelta<sub>20</sub> in pIIa at late cytokinesis in control and *Patronin* RNAi#2-treated cells. iDelta<sub>20</sub> percentage in pIIa is significantly lower upon Patronin depletion (*P* < 0.05; unpaired two-tailed *t*-test). **k, l**, Co-depletion of Klp10A and Patronin suppresses their respective phenotype on endosome asymmetry. **k**, Control, Patronin-, Klp10A-depleted and Patronin/Klp10A-co-depleted SOPs showing mRFP-Pon and iDelta<sub>20</sub> in late cytokinesis. SDCM imaging, *z* projection. **l**, Percentage of iDelta<sub>20</sub> in the pIIa daughter cell at late cytokinesis in control, Patronin-, Klp10A-depleted and Patronin/Klp10A-co-depleted SOPs. Note that the control, *Patronin* RNAi and *Klp10A* RNAi data sets are the same as Fig. 4b, shown here for comparison. Endosome segregation becomes symmetrical upon Patronin or Klp10A depletion, but is restored to control levels upon co-depletion of Patronin and Klp10A (*P* < 0.05; ANOVA followed by Tukey's post hoc test). Note that spindle asymmetry is also suppressed (Extended Data Fig. 6i, j). Unless stated otherwise, *n* corresponds to the number of SOP scored. Scale bars, 5 μm.

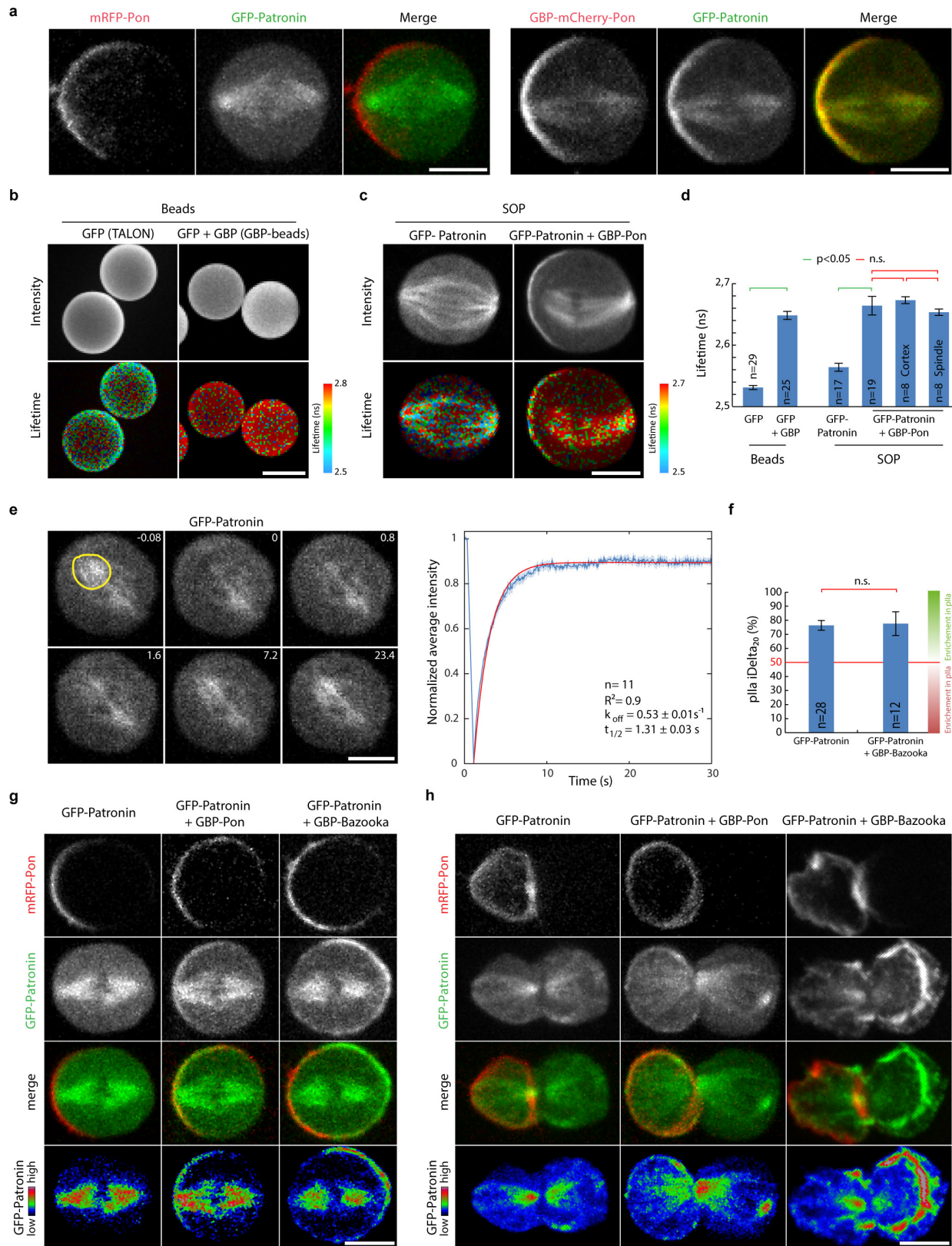


Extended Data Figure 8 | See next page for caption.

**Extended Data Figure 8 | Inversion of central spindle asymmetry by the nanobody assay.** **a**, Model of spindle asymmetry generation by the Patronin/Klp10A machinery. Directly or indirectly, Par complex signalling leads to a biased localization of Patronin, which is thereby enriched on the pIIb side of the central spindle (left panel). Patronin antagonizes Klp10A microtubule depolymerization activity and therefore net microtubule depolymerization activity is stronger on the pIIa side of the spindle (middle panel). As a consequence, the central spindle becomes asymmetric with a higher density in the pIIb side (right panel). **b**, Cartoon of the rationale of the 'nanobody assay' for central spindle inversion (see also 'Rationale of the nanobody experiment' in the Supplementary Discussion). Expression of an anti-GFP nanobody (GBP) fused to an anterior cortex localization domain (Pon; left panel) leads to the specific targeting of GFP-tagged molecules at the anterior cortex. Due to the high affinity of the GBP for GFP<sup>66</sup>, most of the GFP-Patronin molecules are bound to the nanobody (see Extended Data Fig. 9a–d). Owing to the high off-rates of Pon for the cortex<sup>67</sup> and Patronin for the spindle (see Extended Data Fig. 9e) in metaphase, the Pon-GBP/GFP-Patronin complex is in three-state equilibrium between the cytosol, the anterior cortex and the spindle (left panel). In late anaphase, owing to the diffusion barrier at the central spindle (because of crowding), the situation becomes different in pIIa and pIIb: while in pIIb there is a three-state equilibrium (anterior cortex, cytosol, spindle), in pIIa there is only a two-state equilibrium (cytosol, spindle; middle panel). This generates an asymmetry of the amount of Patronin at the spindle: in pIIb there is an additional sink provided by the anterior cortex. As a consequence, the antagonistic activity of Patronin on Klp10a-mediated microtubule depolymerization is depleted in pIIb, thereby biasing microtubule density to higher levels in pIIa (right panel). **c**, Dividing SOPs expressing mRFP-Pon and GFP-Patronin (GFP-Patronin control) or, in addition, expressing GBP-Pon (GFP-Patronin + GBP-Pon)

at late cytokinesis. Like in control cells (upper panel), polarized mRFP-Pon localization is observed upon co-expression of GBP-Pon and GFP-Patronin (middle panel) but, occasionally, low levels of mRFP-Pon are also detected in the pIIa cell cortex (lower panel). Note that the GFP-Patronin control panel is the same as Extended Data Fig. 6g, shown here for comparison. **d**, Frequency of the leaking pIIa mRFP-Pon signal upon expression of the indicated fusion proteins. Cells showing leaking Pon cortical localization in pIIa were excluded from the subsequent analysis of spindle and endosome asymmetries considered in this report. **e–g**, Analysis of spindle inversion in the nanobody assay using a battery of microtubule markers. **e**, Enrichment in pIIb of GFP-Patronin in GFP-Patronin control and GFP-Patronin + GBP-Pon cells. Co-expression of GFP-Patronin and GBP-Pon induces a shift of the frequency distribution of GFP-Patronin asymmetry towards enrichment in pIIa and lower levels of enrichment in pIIb. In these conditions, inverted spindles are observed, which are only very rarely seen in control cells. **f**, Dividing control or GFP-Patronin and GBP-Pon expressing SOPs stained by anti-acetylated- $\alpha$ -tubulin (Ac-tubulin) immunofluorescence. **g**, Enrichment in pIIb of Ac-tubulin in control and GFP-Patronin + GBP-Pon cells. Note the shift towards pIIa enrichment induced by coexpression of GFP-Patronin and GBP-Pon. **h**, Percentage of iDelta<sub>20</sub> in the pIIa daughter cell at late cytokinesis upon expression of the indicated fusion proteins. Co-expression of GBP-Pon with GFP-Patronin causes a statistically significant reduction of iDelta<sub>20</sub> asymmetry compared to control. \* $P < 0.001$  compared to control, as calculated by ANOVA followed by Tukey's post hoc test. Comparisons between the other genetic backgrounds and the control or between themselves are not significant. All images in this figure correspond to SDCM imaging (maximum-intensity projection). n, number of SOPs scored. Scale bars, 5  $\mu$ m.

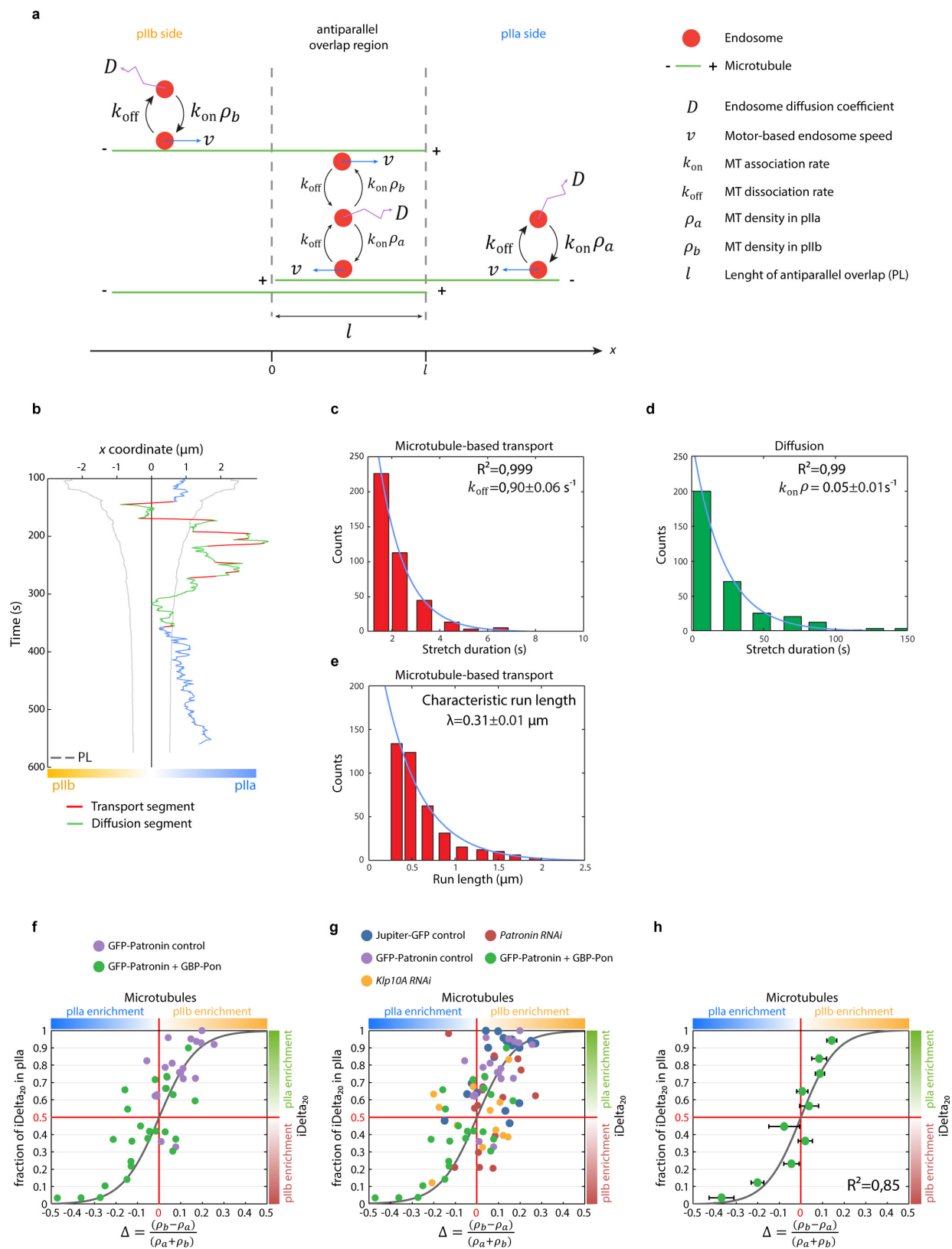




Extended Data Figure 9 | See next page for caption.

**Extended Data Figure 9 | Rationale of the nanobody assay. a–d,** Most GFP–Patronin in the SOP is bound to GBP–Pon. **a,** Dividing SOPs expressing mRFP–Pon and GFP–Patronin (left panels) or GBP–mCherry–Pon and GFP–Patronin (right panels) were imaged by SDCM (maximum-intensity projection). Note that GBP–mCherry–Pon colocalizes with GFP–Patronin both at the cortex and at the spindle. **b,** GBP binding increases the fluorescence lifetime of GFP *in vitro*. Purified (His)<sub>6</sub>–GFP was coupled to TALON beads through the (His)<sub>6</sub> tag (left panels) or to GBP beads through GFP itself (right panels). The GFP fluorescence lifetime was imaged by FLIM (see Methods). Bottom panels correspond to images where abundance of GFP (pixel intensity) is combined with (multiplied by) the lifetime of GFP (pixel colour). Note that lifetime increases (that is, more red pixels) in the GBP beads when compared to the TALON beads (more cyan pixels). **c,** GBP binding increases the fluorescence lifetime of GFP–Patronin *in vivo*. Dividing SOPs expressing mRFP–Pon and GFP–Patronin (GFP–Patronin, control) or expressing GBP–Pon in addition (GFP–Patronin + GBP–Pon) at metaphase imaged by FLIM. As in **b**, bottom panel corresponds to images where abundance of GFP–Patronin (pixel intensity) is combined with its lifetime (pixel colour). Note that lifetime increases (that is, more red pixels) in the GBP–Pon sample. **d,** GFP lifetime measurements in the samples described in **b** and **c** (mean  $\pm$  s.e.m. of different beads or cells; n, number of beads or cells scored). Note that lifetimes of GFP–Patronin are the same in the spindle and the cortex in the GFP–Patronin + GBP–Pon SOP, and that this homogenous lifetime value in the SOP corresponds to that of (His)<sub>6</sub>–GFP bound to nanobody beads. This suggests that most GFP–Patronin molecules are bound to the nanobody in the SOP. Please note that in the experiments described in **b** (purified GFP) and in **c** (GFP–Patronin fusion in the fly) the term GFP refers to the enhanced GFP variant (eGFP). **e,** Fluorescence recovery after

photobleaching (FRAP) of GFP–Patronin at the spindle. Left panel, images from the FRAP experiment. Fluorescence of GFP–Patronin at the mitotic spindle was bleached (yellow region of interest) and fluorescence recovery was monitored over time. Elapsed time, indicated in seconds. Right panel, average recovery of 11 cells (blue curve; mean  $\pm$  s.e.m.) was fitted to a single exponential model (red line), providing an estimate of the half-life of GFP–Patronin on microtubules of  $1.31 \pm 0.03$  s (95% confidence interval). This indicates that Patronin exchanges rapidly between the cytosol and the spindle. **f–h,** Co-expression of GBP–Bazooka and GFP–Patronin does not cause central spindle inversion in the SOP. **f,** Percentage of iDelta<sub>20</sub> in the pIIa daughter cell at late cytokinesis in SOPs expressing GFP–Patronin (GFP–Patronin) or expressing GBP–Bazooka in addition (GFP–Patronin + GBP–Bazooka). Co-expression of GBP–Bazooka with GFP–Patronin does not cause a statistically significant reduction of iDelta<sub>20</sub> asymmetry compared to control. ( $P = 0.847$ , unpaired two-tailed *t*-test). **g,** Dividing SOPs expressing mRFP–Pon and eGFP–Patronin (GFP–Patronin) or expressing GBP–Pon in addition (GFP–Patronin + GBP–Pon) or GBP–Bazooka in addition (GFP–Patronin + GBP–Bazooka) in metaphase. SDCM imaging (single plane). Bottom panels correspond to the rainbow lookup table. GBP–Pon targets GFP–Patronin asymmetrically to the cortex of the pIIb and, conversely, GBP–Bazooka targets GFP–Patronin asymmetrically to the cortex of the pIIa. Note that, in metaphase, the mitotic spindle is symmetric in control, but also in the two GBP conditions. **h,** Dividing SOPs as in **g** in late cytokinesis. GBP–Pon inverts the asymmetry of the central spindle, whereas GBP–Bazooka does not (it shows the same asymmetry as in control). SDCM imaging (maximum-intensity projection, except GFP–Patronin in the GFP–Patronin + GBP–Bazooka image, which corresponds to a single plane). n, number of SOPs scored. Scale bars, 5  $\mu$ m (**a**, **c**, **e**, **g**, **h**) and 40  $\mu$ m (**b**).



Extended Data Figure 10 | See next page for caption.



**Extended Data Figure 10 | Model of motility onto an asymmetric central spindle.** **a**, Schematic representation of a model of bidirectional motor transport, which generates a biased spatial distribution of endosomes through an antiparallel microtubule overlap (see Supplementary Information equations). **b–e**, Determination of  $k_{\text{off}}$  and  $k_{\text{on}}$  of endosomes for microtubules and transport run length in control cells (see also ‘Determination of  $k_{\text{off}}$  and  $k_{\text{on}}$  and transport run length’ in Methods for details). **b**, Representative endosome track at the central spindle (along the  $x$  axis, as in Fig. 2e) after automated detection of transport segments (that is, motility occurs by microtubule-based transport; red) and of diffusion segments (that is, movement occurs by diffusion; green). Blue segments correspond to the initial and final segments of the track, the total length of which cannot be determined and are therefore excluded from the analysis. Grey line, extent of the Pavarotti antiparallel overlap of microtubules as in Fig. 2. **c–e**, Distribution of the duration (**c**, **d**) and the run length (**e**) of transport segments (**c**, **e**) and/or diffusion-segments (**d**) from 101 tracks. Exponential fits (blue lines) give  $k_{\text{off}} = 0.90 \pm 0.06 \text{ s}^{-1}$  (**c**) and  $k_{\text{on}} = 0.05 \pm 0.01 \text{ s}^{-1}$  (**d**; 95% confidence interval). Characteristic run length was not estimated from exponential fit of this distribution in **e**, but instead by the method of Thorn and Vale (ref. 65), which relies on the fitting of the cumulative distribution of the run length (see Methods). The estimated characteristic transport run

length is  $\lambda = 0.31 \pm 0.01 \mu\text{m}$  ( $R^2 = 0.98$ ; 95% confidence interval).

**f**, Endosome asymmetric segregation (that is,  $\frac{P_{\text{pIIa}}}{P_{\text{pIIa}} + P_{\text{pIIb}}}$ ;  $P_{\text{pIIa}}$ , iDelta<sub>20</sub> signal in pIIa;  $P_{\text{pIIb}}$ , in pIIb) as a function of the central spindle asymmetry (that is,  $\Delta = \frac{\rho_{\text{b}} - \rho_{\text{a}}}{\rho_{\text{b}} + \rho_{\text{a}}}$ ;  $\rho_{\text{b}}$ , GFP–Patronin density in the pIIb side;  $\rho_{\text{a}}$ , in pIIa) for GFP–Patronin control ( $n = 21$ ) or GFP–Patronin + GBP–Pon ( $n = 25$ ) cells at during late cytokinesis ( $\sim 600 \text{ s}$  in registered time). The same data are shown in bins in Fig. 4h. The experimental data points fall on the theoretical curve (grey line corresponding to equation (36), with measured parameters values; see Supplementary Information equations) which accounts for the different levels of spindle reversal and their corresponding endosome segregation reversal measured in this assay. **g**, Fraction of iDelta<sub>20</sub> versus microtubule enrichment as in **f**, considering the data from all the genotypes presented in this study. The model (grey line, as in **f**) for asymmetric bidirectional transport accounts for all spindle asymmetry/endosome asymmetries found across all the backgrounds presented in this study. **h**, Fraction of iDelta<sub>20</sub> versus microtubule enrichment as in **g**. The model outside of the limit case (grey line corresponding to equation (37) with measured parameters values and fitted  $l = 3.2 \mu\text{m}$ , see Supplementary Information equations) also accounts for all spindle asymmetry/endosome asymmetries found across all the backgrounds presented in this study.

# Replication stress activates DNA repair synthesis in mitosis

Sheroy Minocherhomji<sup>1,\*</sup>, Songmin Ying<sup>2,3\*</sup>, Victoria A. Bjerregaard<sup>1</sup>, Sara Bursomanno<sup>1</sup>, Aiste Aleliunaite<sup>1</sup>, Wei Wu<sup>1</sup>, Hocine W. Mankouri<sup>1</sup>, Huahao Shen<sup>2,4</sup>, Ying Liu<sup>1</sup> & Ian D. Hickson<sup>1</sup>

**Oncogene-induced DNA replication stress has been implicated as a driver of tumorigenesis<sup>1</sup>. Many chromosomal rearrangements characteristic of human cancers originate from specific regions of the genome called common fragile sites (CFSs)<sup>2–5</sup>. CFSs are difficult-to-replicate loci that manifest as gaps or breaks on metaphase chromosomes (termed CFS ‘expression’), particularly when cells have been exposed to replicative stress<sup>6</sup>. The MUS81–EME1 structure-specific endonuclease promotes the appearance of chromosome gaps or breaks at CFSs following replicative stress<sup>7–9</sup>. Here we show that entry of cells into mitotic prophase triggers the recruitment of MUS81 to CFSs. The nuclease activity of MUS81 then promotes POLD3-dependent DNA synthesis at CFSs, which serves to minimize chromosome mis-segregation and non-disjunction. We propose that the attempted condensation of incompletely duplicated loci in early mitosis serves as the trigger for completion of DNA replication at CFS loci in human cells. Given that this POLD3-dependent mitotic DNA synthesis is enhanced in aneuploid cancer cells that exhibit intrinsically high levels of chromosomal instability (CIN<sup>+</sup>) and replicative stress, we suggest that targeting this pathway could represent a new therapeutic approach.**

CFSs normally replicate late in S phase and this can be delayed even further following replicative stress<sup>10</sup>. To define when DNA synthesis occurs at CFSs, we treated U2OS cells with replicative stress in the form of a low dose (0.4  $\mu$ M) of the DNA polymerase inhibitor aphidicolin (APH), and then added 5-ethynyl-2'-deoxyuridine (EdU) for 30 min to visualize new DNA synthesis (Extended Data Fig. 1a). In addition to scheduled DNA replication in S phase, over 40% of the mitotic cells, including those in anaphase, contained EdU foci (Fig. 1a, b and Extended Data Fig. 1b, c). Given that the combined length of the initial mitotic phases is >30 min<sup>11</sup> (Extended Data Fig. 1d and Supplementary Video 1), these data suggested that some DNA synthesis might occur in early mitosis. To investigate this, cells were pre-treated with/without low-dose APH, arrested in late G2 using the CDK1 inhibitor RO3306 (ref. 12), and then released into mitosis in the presence of EdU (Extended Data Figs 1e–g and 2a–d). EdU incorporation was only detectable in APH-treated cells that had initiated mitosis (Extended Data Fig. 2e, f). If EdU was added to nocodazole-arrested cells in prometaphase, EdU foci were not observed (Extended Data Fig. 2e–g). These data indicate that replicative-stress-induced DNA synthesis can occur after entry into mitosis, but only before the nocodazole arrest point in prometaphase<sup>13</sup>. Approximately 80% of the EdU foci in condensed nuclei co-localized with FANCD2 twin foci (Fig. 1c, d), which define the location of CFSs<sup>14</sup>. Moreover, most (80%) 4',6-diamidino-2-phenylindole (DAPI)-negative gaps/breaks or defined constrictions in condensed chromosomes were positive for EdU (Fig. 1e, f and Extended Data Fig. 3a, b), as were the FRA3B and FRA16D CFS loci (Extended Data Fig. 3c, d). In contrast, mitotic EdU

foci rarely co-localized with centromeres or telomeres (Extended Data Fig. 3e, f). Hence, mitotic DNA synthesis occurs predominantly at CFSs.

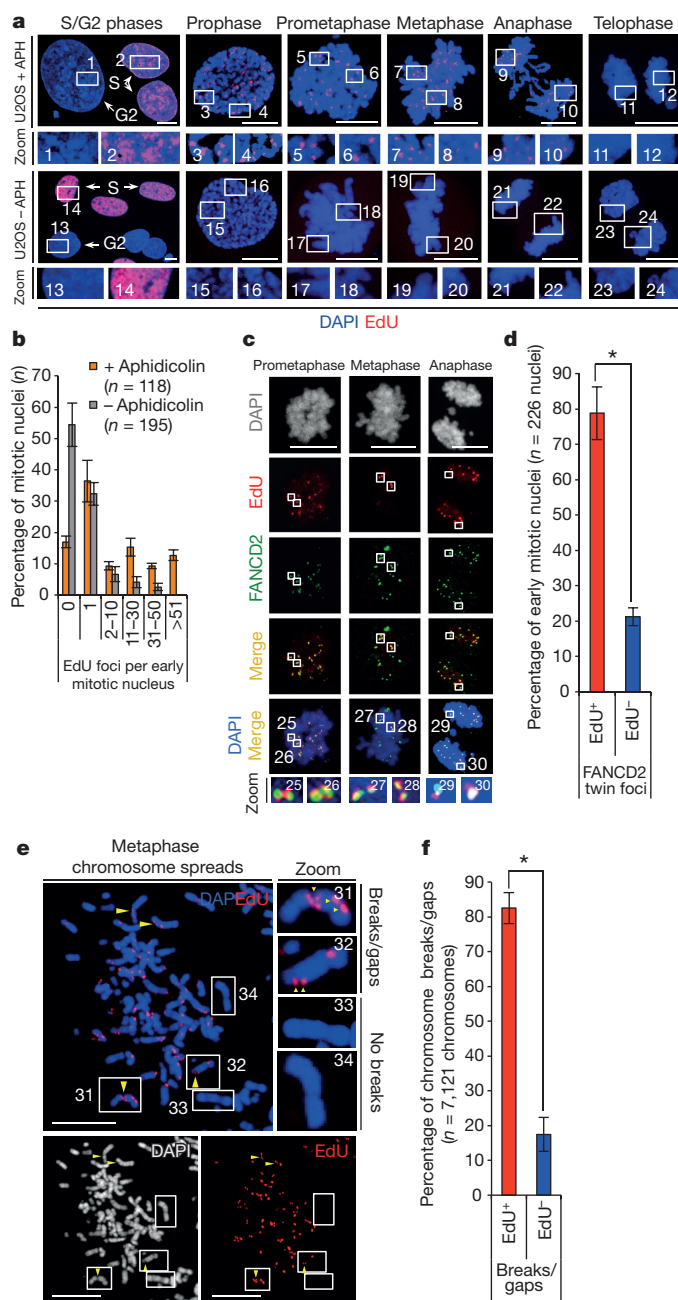
Consistent with a role for the MUS81–EME1 complex in promoting CFS expression<sup>7–9</sup>, depletion of either MUS81 or EME1 inhibited EdU incorporation in mitotic cells (Fig. 2a–d and Extended Data Fig. 4a). A short interfering RNA (siRNA)-resistant *MUS81* cDNA, but not a catalytically inactive version (*MUS81(D338A/D339A)*) (refs 15, 16; Extended Data Fig. 4a), could restore mitotic EdU incorporation and CFS expression in *MUS81* siRNA-depleted cells (Fig. 2a–e). Therefore, MUS81 endonuclease activity is required to promote DNA synthesis in mitosis. SLX4, a scaffold for various nucleases including MUS81–EME1 (refs 17, 18), also co-localized with CFSs throughout mitosis (Extended Data Fig. 4b–d). This localization did not require MUS81. Depletion of SLX4 (Extended Data Fig. 4a) also led to: (1) reduced EdU incorporation at CFSs in mitosis; (2) reduced CFS expression; (3) defective recruitment of MUS81 to CFSs; (4) increased FANCD2-associated ultra-fine anaphase DNA bridges (UFBs); and (5) increased 53BP1 nuclear bodies that form around unprocessed CFSs<sup>7,19,20</sup> (Fig. 2a–g and Extended Data Figs 4e–h and 5a–c).

To study whether the key events occurring in prophase were required for EdU incorporation, we depleted SMC2, which is required for chromosome condensation<sup>21</sup>, and disrupted the pathway required for the release of sister-chromatid arm cohesion by either depleting WAPL (Wings apart protein-like)<sup>22</sup> or inhibiting PLK1 (PLK1i) in early mitosis using BI-2536 (ref. 23). Depletion of either SMC2 or WAPL led to characteristic alterations in chromosome morphology and compaction (Extended Data Fig. 5d–f), but failed to activate interphase checkpoints (Fig. 2a, b and Extended Data Fig. 4a) or generate increased numbers of DNA damage foci in S phase compared to control-siRNA-depleted cells (Extended Data Fig. 5g, h). However, SMC2 or WAPL depletion did lead to a failure to recruit MUS81 to chromatin, decreased CFS expression, and an abolition of EdU incorporation at CFSs in mitosis (Fig. 2b–g). Furthermore, these cells exhibited increased numbers of anaphase bridges (Extended Data Fig. 4e–h) and 53BP1 nuclear bodies (Extended Data Fig. 5a–c). Co-localization of SLX4 with FANCD2 twin foci in mitosis was unaffected (Extended Data Fig. 4d). Notably, PLK1i in mitosis also inhibited mitotic EdU incorporation (Extended Data Fig. 5i, j), implying that CFS-associated DNA synthesis occurs after, or concomitant with, release of sister chromatid arm cohesion.

To examine the DNA polymerase(s) responsible for mitotic DNA synthesis, we released G2-arrested cells pre-treated with/without low-dose APH into mitosis in the presence of a high dose of APH (2  $\mu$ M). These cells failed to incorporate EdU (Fig. 3a–c), and demonstrated decreased CFS expression (Extended Data Fig. 6a, b). Increased numbers of 53BP1 nuclear bodies were evident in G1 cells (Fig. 3d, e), which rarely co-localized with EdU that was incorporated during the previous mitosis (Extended Data Fig. 6c–f),

<sup>1</sup>Center for Chromosome Stability and Center for Healthy Aging, Department of Cellular and Molecular Medicine, University of Copenhagen, Panum Institute, Blegdamsvej 3B, 2200 Copenhagen N, Denmark. <sup>2</sup>Department of Respiratory and Critical Care Medicine of the Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310009, China. <sup>3</sup>Department of Pharmacology, Zhejiang University School of Medicine, Hangzhou 310058, China. <sup>4</sup>State Key Laboratory of Respiratory Disease (SKLRD), Guangzhou 510120, China.

\*These authors contributed equally to this work.



**Figure 1 | Replicative stress activates DNA synthesis at CFSs in mitosis.**

**a, b**, DNA synthesis (EdU foci, red) in intact or condensed mitotic nuclei (DAPI, blue). **c, d**, Co-localized EdU foci (red) and FANCD2 twin foci (green) in mitotic nuclei (DAPI, blue). **e, f**, EdU incorporation (arrows) at DAPI-negative breaks/gaps on metaphase chromosomes (DAPI, blue). Panels show representative images (**a, c, e**) and quantification (**b, d, f**). Selected regions are magnified in numbered panels. Scale bars, 10  $\mu$ m. Data are means of three independent experiments. Error bars indicate s.e.m. \* $P < 0.01$ .

suggesting that a failure to perform mitotic DNA synthesis adversely impacts daughter cells. To examine this, we analysed the fate of FRA3B and FRA16D loci in two cell lines (HCT-116 and MRC5) with a largely diploid karyotype (Extended Data Fig. 6g). A high dose of APH in mitosis generated an increased frequency of CFS-associated UFBs, increased non-disjunction of chromosomes 3 and 16 harbouring FRA3B and FRA16D, and reduced cell survival (Fig. 3f–h and Extended Data Fig. 6h–l). Therefore, mitotic DNA synthesis is required to counteract potentially lethal chromosome mis-segregation and non-disjunction.

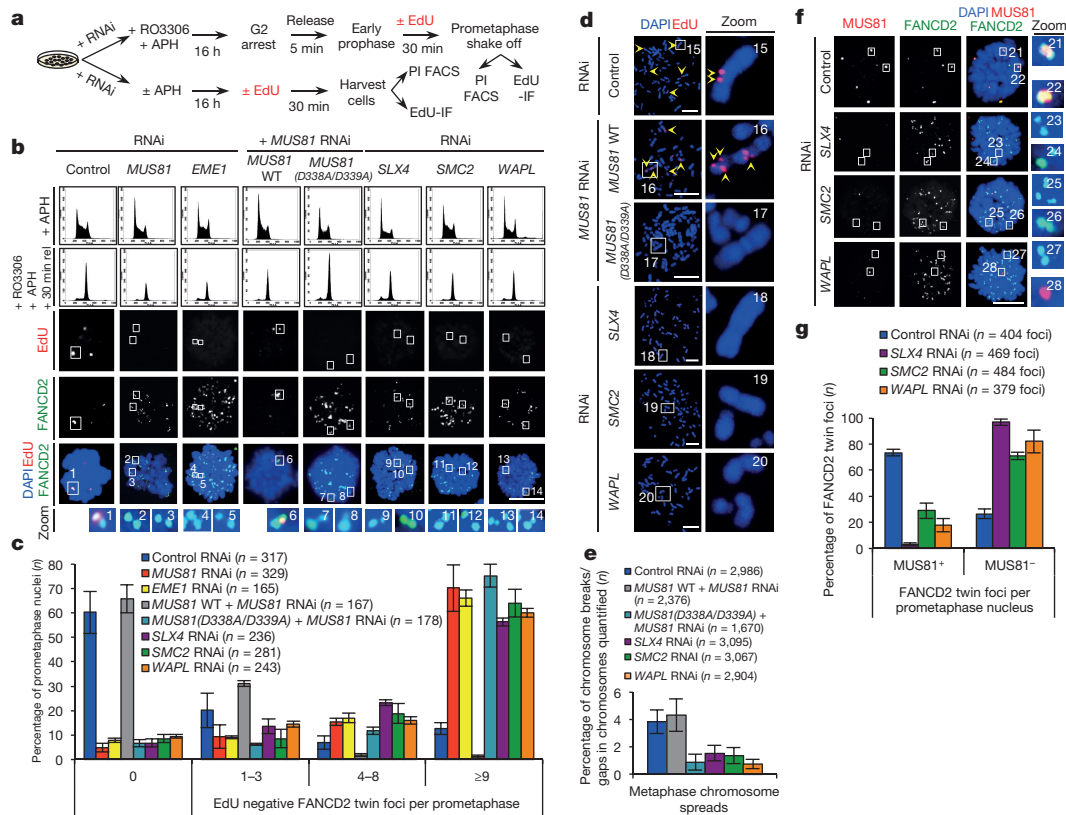
Subcellular fractionation analysis revealed that PCNA, the MCM complex, MUS81–EME1 and POLD1 and POLD3 (catalytic and non-catalytic subunits, respectively, of DNA polymerase  $\delta$ ) were all associated with mitotic chromatin after replicative stress (Fig. 4a, b and Extended Data Fig. 7a, b). The presence of the MCM complex on mitotic chromatin did not require MUS81 (Extended Data Fig. 4a). Notably, MUS81 was largely absent from chromatin in the G2 phase (Fig. 4a, b), and was again recruited to chromatin after entry into mitosis at the same time as EME1 was phosphorylated<sup>17</sup>. POLD3 formed APH-inducible foci in early mitosis, but very rarely in anaphase (Fig. 4c, d), similar to the pattern reported previously for MUS81 (ref. 7). POLD3, like its yeast orthologue Pol32, is required for a specialized form of DNA repair termed break-induced replication (BIR), which is POLD4-independent<sup>24–26</sup>. Depletion of POLD3 or POLD4 did not elicit a checkpoint response or an increase in the DNA damage response (Extended Data Fig. 7c–i), and did not prevent MUS81 recruitment to CFSs in mitosis (Fig. 4e, f). However, POLD3 (but not POLD4) depletion prevented mitotic DNA synthesis at CFSs (Fig. 4g, h and Extended Data Fig. 8a), and led to increased levels of DNA strand breaks (TUNEL<sup>+</sup>) on metaphase chromosomes (Fig. 4i, j). Expression of siRNA-resistant POLD3 restored mitotic DNA synthesis in POLD3-depleted cells (Fig. 4g, h and Extended Data Fig. 8b). Moreover, APH-induced CFS expression, and suppression of CFS-associated UFBs and 53BP1 bodies, required POLD3 but not POLD4 (Fig. 4k, l and Extended Data Fig. 8c–h). We propose that metaphase chromosome breaks/gaps are associated with *de novo* mitotic DNA synthesis promoted by the POLD3-associated DNA polymerase  $\delta$  complex.

We next examined the relationship between POLD3 and MUS81 in regulating recruitment of DNA replication/repair factors to CFSs in mitosis. Depletion of POLD3 or POLD4 had negligible effects on the association of MUS81, POLD1 or PCNA with S phase or mitotic chromatin (Extended Data Fig. 8i, j). Depletion of MUS81, however, reduced POLD3 and POLD4 (but not POLD1 or PCNA) recruitment in S phase, suggesting a potential role for MUS81, POLD3 and POLD4 in DNA repair after APH-induced replication stress. MUS81 depletion also abrogated recruitment of POLD3 and POLD4 to chromatin in prometaphase cells, but recruitment of POLD1 and PCNA was also impaired (Extended Data Fig. 8i, k).

Because the structure and expression of the *POLD3* gene is altered in some cancers<sup>27</sup>, we investigated whether DNA synthesis in early mitosis was a specific feature of aneuploid/CIN<sup>+</sup> human cancer cell lines. Low-dose (0.4  $\mu$ M) APH elicited FANCD2 DNA damage focus formation in all ten cell lines studied, irrespective of their ploidy or whether they were cancer-derived, untransformed or virally transformed (Extended Data Figs 9a and 10a–c). Co-localization of EdU foci and CFSs in mitosis was evident in all cell lines examined (Extended Data Fig. 10d–g), indicating that mitotic DNA synthesis is not limited to cancer cells; however, the frequency of co-localized EdU foci and FANCD2 twin foci in mitosis was generally higher in the aneuploid/CIN<sup>+</sup> cancer cell lines (Extended Data Fig. 10f, g).

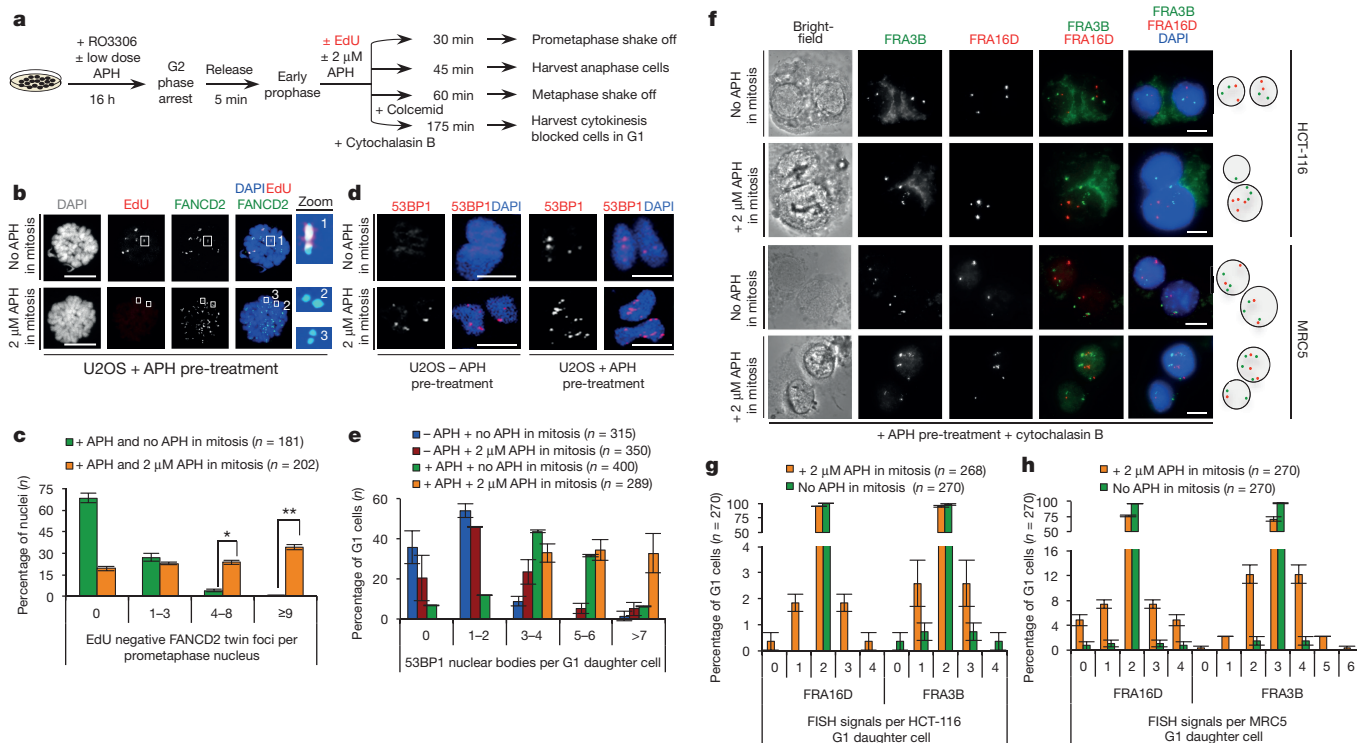
We propose a model (Extended Data Fig. 9b, c) whereby cleavage of stalled replication forks at CFSs in early mitosis by SLX4-associated MUS81–EME1 promotes POLD3-dependent DNA synthesis at these loci. Recent data suggest that an analogous TOPBP1-dependent process also exists in chicken DT40 cells<sup>28</sup>. We suggest that the mitotic DNA synthesis described here might occur via BIR, which produces a ‘migrating bubble’ that drives conservative DNA synthesis<sup>25,29,30</sup>, although whether the mechanism of mitotic DNA synthesis in humans is identical to that shown for DSB-induced BIR in yeast is unclear. Because the DNA synthesis reported here occurs after the initiation of DNA condensation in prophase, we speculate that the mechanical forces of chromosome condensation could expose under-replicated regions (that remain ‘primed’ for replication) at CFSs, and subsequently promote the recruitment of the



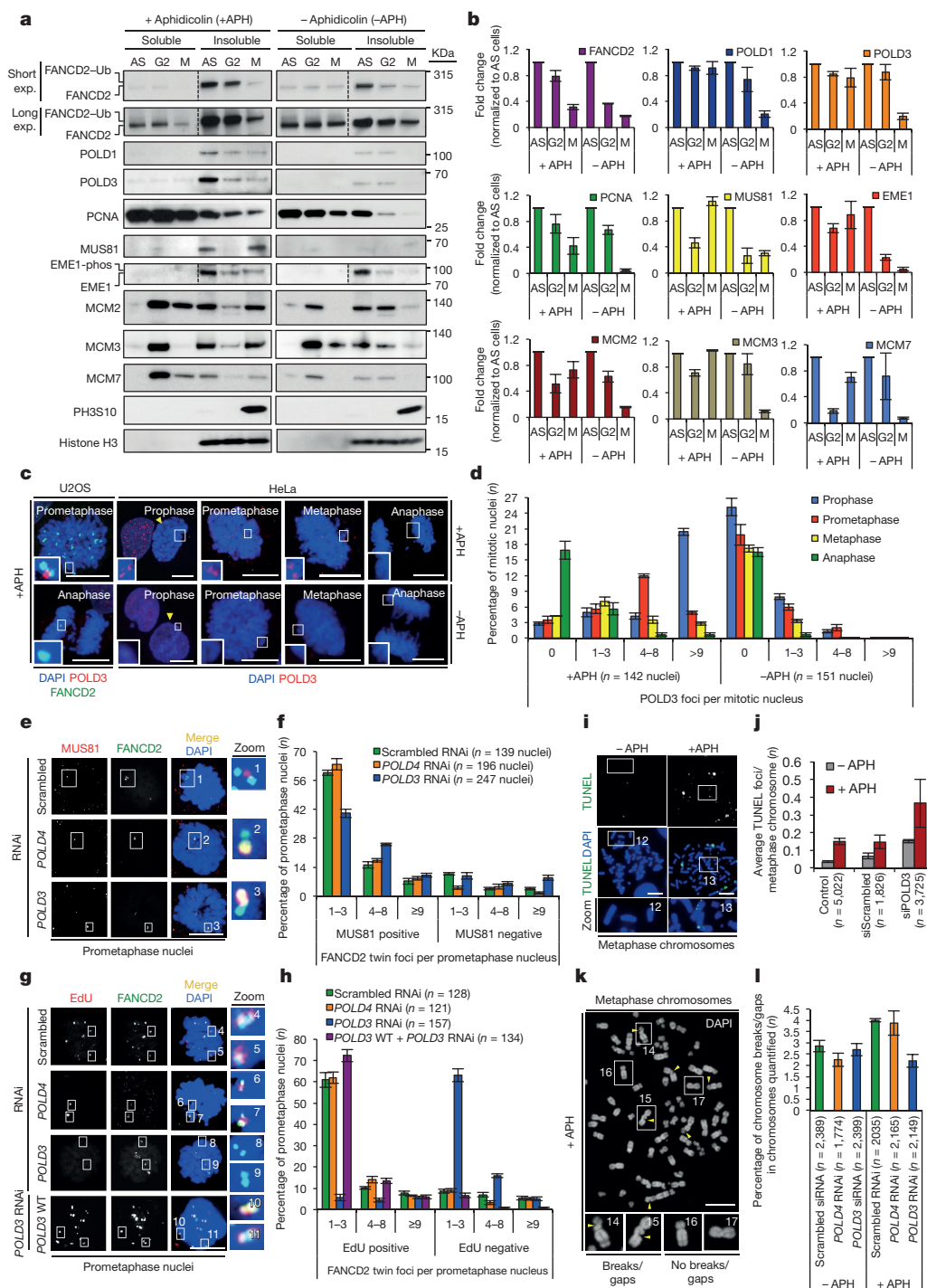


**Figure 2 | MUS81, SLX4, SMC2 and WAPL promote DNA synthesis at CFSs in mitosis.** **a**, Experimental workflow. **b**, **c**, FACS traces for asynchronous and prometaphase cells (top two panels), representative images (bottom three panels) and quantification (**c**) of co-localized EdU (red) and FANCD2 twin foci (green). **d**, **e**, Representative images (**d**) and quantification (**e**) of EdU foci (red) at DAPI-negative chromosome

breaks/gaps. **f**, **g**, Representative images (**f**) and quantification (**g**) of co-localized MUS81 (red) and FANCD2 twin foci (green). DNA was stained using DAPI (blue). Selected regions are magnified and numbered. Scale bars, 10  $\mu$ m. Data are means of three independent experiments. Error bars indicate s.e.m.



**Figure 3 | DNA synthesis at CFSs in mitosis suppresses non-disjunction.** **a**, Experimental workflow. **b**, **c**, DNA synthesis (EdU, red) and FANCD2 twin foci (green) in prometaphase cells. \* $P < 0.01$ , \*\* $P < 0.001$ . **d**, **e**, 53BP1 nuclear bodies (red) in early G1 daughter cells. **f**–**h**, FISH analysis; DNA was stained using DAPI (blue). Panels show representative images (**b**, **d**, **f**) and quantification (**c**, **e**, **g**, **h**). Selected regions are magnified in numbered panels. Scale bars, 10  $\mu$ m. Data are means of at least three independent experiments. Error bars indicate s.e.m.



**Figure 4 | POLD3 promotes DNA synthesis in mitosis. a, b**, Western blots. Loading control, H3. PH3S10 indicates mitotic chromatin. **c, d**, Co-localized POLD3 (red) and FANCD2 twin foci (green) in mitosis. **e–h**, Co-localized MUS81 (red) or EdU (red) foci at FANCD2 twin foci

(green). **i–l**, TUNEL foci (green) or DAPI-negative chromosome breaks/gaps (arrows); TUNEL was stained using DAPI (blue). Panels show representative images (**a, c, e, g, i, k**) and quantification (**b, d, f, h, j, l**). Scale bars, 10 μm. Data are means of three independent experiments. Error bars indicate s.e.m.

SLX4-associated MUS81–EME1 endonuclease. Our findings reveal that mitotic DNA synthesis is detectable at sites of DAPI-negative chromosomal regions, suggesting that these cytogenetically defined loci comprise DNA synthesized in early mitosis, rather than DNA strand breaks. It is conceivable that this POLD3-dependent pathway used to counteract constitutively high levels of replicative stress might represent a novel anticancer drug target.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 23 January; accepted 8 October 2015.

Published online 2 December 2015.

- Halazonetis, T. D., Gorgoulis, V. G. & Bartek, J. An oncogene-induced DNA damage model for cancer development. *Science* **319**, 1352–1355 (2008).
- Burrow, A. A., Williams, L. E., Pierce, L. C. & Wang, Y. H. Over half of breakpoints in gene pairs involved in cancer-specific recurrent translocations are mapped to human chromosomal fragile sites. *BMC Genomics* **10**, 59 (2009).
- Arlt, M. F., Durkin, S. G., Ragland, R. L. & Glover, T. W. Common fragile sites as targets for chromosome rearrangements. *DNA Repair (Amst.)* **5**, 1126–1135 (2006).

4. Le Tallec, B. *et al.* Common fragile site profiling in epithelial and erythroid cells reveals that most recurrent cancer deletions lie in fragile sites hosting large genes. *Cell Reports* **4**, 420–428 (2013).
5. Burrell, R. A. *et al.* Replication stress links structural and numerical cancer chromosomal instability. *Nature* **494**, 492–496 (2013).
6. Glover, T. W., Berger, C., Coyle, J. & Echo, B. DNA polymerase  $\alpha$  inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. *Hum. Genet.* **67**, 136–142 (1984).
7. Ying, S. *et al.* MUS81 promotes common fragile site expression. *Nature Cell Biol.* **15**, 1001–1007 (2013).
8. Naim, V., Wilhelm, T., Debatisse, M. & Rosselli, F. ERCC1 and MUS81–EME1 promote sister chromatid separation by processing late replication intermediates at common fragile sites during mitosis. *Nature Cell Biol.* **15**, 1008–1015 (2013).
9. Minocherhomji, S. & Hickson, I. D. Structure-specific endonucleases: guardians of fragile site stability. *Trends Cell Biol.* **24**, 321–327 (2014).
10. Le Beau, M. M. *et al.* Replication of a common fragile site, FRA3B, occurs late in S phase and is delayed further upon induction: implications for the mechanism of fragile site induction. *Hum. Mol. Genet.* **7**, 755–761 (1998).
11. Meraldi, P., Draviam, V. M. & Sorger, P. K. Timing and checkpoints in the regulation of mitotic progression. *Dev. Cell* **7**, 45–60 (2004).
12. Vassilev, L. T. *et al.* Selective small-molecule inhibitor reveals critical mitotic functions of human CDK1. *Proc. Natl Acad. Sci. USA* **103**, 10660–10665 (2006).
13. Beaudouin, J., Gerlich, D., Daigle, N., Eils, R. & Ellenberg, J. Nuclear envelope breakdown proceeds by microtubule-induced tearing of the lamina. *Cell* **108**, 83–96 (2002).
14. Chan, K. L., Palmai-Pallag, T., Ying, S. & Hickson, I. D. Replication stress induces sister-chromatid bridging at fragile site loci in mitosis. *Nature Cell Biol.* **11**, 753–760 (2009).
15. Hanada, K. *et al.* The structure-specific endonuclease Mus81 contributes to replication restart by generating double-strand DNA breaks. *Nature Struct. Mol. Biol.* **14**, 1096–1104 (2007).
16. Ehmsen, K. T. & Heyer, W.-D. *Saccharomyces cerevisiae* Mus81–Mms4 is a catalytic, DNA structure-selective endonuclease. *Nucleic Acids Res.* **36**, 2182–2195 (2008).
17. Wyatt, H. D., Sarbajna, S., Matos, J. & West, S. C. Coordinated actions of SLX1–SLX4 and MUS81–EME1 for Holliday junction resolution in human cells. *Mol. Cell* **52**, 234–247 (2013).
18. Castor, D. *et al.* Cooperative control of Holliday junction resolution and DNA repair by the SLX1 and MUS81–EME1 nucleases. *Mol. Cell* **52**, 221–233 (2013).
19. Lukas, C. *et al.* 53BP1 nuclear bodies form around DNA lesions generated by mitotic transmission of chromosomes under replication stress. *Nature Cell Biol.* **13**, 243–253 (2011).
20. Harrigan, J. A. *et al.* Replication stress induces 53BP1-containing OPT domains in G1 cells. *J. Cell Biol.* **193**, 97–108 (2011).
21. Hirano, T. Condensins: universal organizers of chromosomes with diverse functions. *Genes Dev.* **26**, 1659–1678 (2012).
22. Tedeschi, A. *et al.* Wapl is an essential regulator of chromatin structure and chromosome segregation. *Nature* **501**, 564–568 (2013).
23. Lénárt, P. *et al.* The small-molecule inhibitor BI 2536 reveals novel insights into mitotic roles of polo-like kinase 1. *Curr. Biol.* **17**, 304–315 (2007).
24. Costantino, L. *et al.* Break-induced replication repair of damaged forks induces genomic duplications in human cells. *Science* **343**, 88–91 (2014).
25. Donnianni, R. A. & Symington, L. S. Break-induced replication occurs by conservative DNA synthesis. *Proc. Natl Acad. Sci. USA* **110**, 13475–13480 (2013).
26. Lydeard, J. R., Jain, S., Yamaguchi, M. & Haber, J. E. Break-induced replication and telomerase-independent telomere maintenance require Pol32. *Nature* **448**, 820–823 (2007).
27. Dunlop, M. G. *et al.* Common variation near CDKN1A, POLD3 and SHROOM2 influences colorectal cancer risk. *Nature Genet.* **44**, 770–776 (2012).
28. Pedersen, R. T., Kruse, T., Nilsson, J., Oestergaard, V. H. & Lisby, M. TopBP1 is required at mitosis to reduce transmission of DNA damage to G1 daughter cells. *J. Cell Biol.* **210**, 565–582 (2015).
29. Wilson, M. A. *et al.* Pif1 helicase and Pol $\delta$  promote recombination-coupled DNA synthesis via bubble migration. *Nature* **502**, 393–396 (2013).
30. Saini, N. *et al.* Migrating bubble during break-induced replication drives conservative DNA synthesis. *Nature* **502**, 389–392 (2013).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank T. Palmai-Pallag, L. Che, B. Tian, M. Xing, X. Geng and T. Wass for technical help. This study was supported by The European Research Council, Danish National Research Foundation and Nordea Foundation (H.W.M., Y.L. and I.D.H.), the National 1000 Talents Program, the National Natural Science Foundation of China (31370901, 81422031), Zhejiang Provincial Natural Science Foundation of China (LR14H160001) (S.Y.), National Key Scientific and Technology Support Program: Collaborative innovation of Clinical Research for chronic obstructive pulmonary disease and lung cancer, no. 2013BAI09B09 (H.S. and S.Y.), and The Danish Medical Research Council (S.M.). I.D.H. is a Qiushi Guest Professor at Zhejiang University.

**Author Contributions** S.M., S.Y., V.A.B., S.B., A.A. and W.W. carried out experiments. S.M., S.Y., H.W.M., H.S., Y.L. and I.D.H. designed experiments and interpreted results. S.M., H.W.M., Y.L. and I.D.H. wrote the manuscript, and all authors edited it.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.S. (huahaoshen@163.com) or Y.L. (ying@sund.ku.dk) or I.D.H. (iandh@sund.ku.dk).



## METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

**Cell culture.** Osteosarcoma cells (U2OS), cervical cancer cells (HeLa), colorectal cancer cells (HT-29, CC20 and DLD1) and non-cancer-derived HS68 untransformed foreskin fibroblasts were maintained in Dulbecco's modified Eagle's medium (DMEM). The colorectal cancer cell lines HCT-116 and LS411 were maintained in McCoy's Essential medium and RPMI 1640 medium, respectively. GM00637 SV40-transformed skin fibroblasts and MRC5 untransformed lung fibroblasts were maintained in Eagle's Minimum Essential medium (MEM). In all cases, the medium was supplemented with 10% fetal bovine serum (FBS) and antibiotics. All cells were maintained at 37°C in an atmosphere of 5% CO<sub>2</sub>. The cell lines studied were obtained from the ATCC, with the exception of CC20 and LS411, which were provided by W. F. Bodmer (University of Oxford, UK), and are not among those listed as commonly misidentified by the International Cell Line Authentication Committee. All cell lines were subjected to monthly mycoplasma testing (using MycoAlert; Lonza Group Ltd) and found to be negative. The identity of the cell lines was validated by STR profiling (ATCC) and by analysis of chromosome number in metaphase spreads.

**RNA interference.** MUS81 was depleted using either of three si/shRNAs designated MUS81 shRNA 1, MUS81 shRNA 2 (ref. 7) or MUS81 siRNA 3 (Cat 130845, Ambion). POLD3 was depleted using an equimolar mix of four pooled siRNAs (L-026692-01-0010; Dharmacon) as described previously<sup>24</sup>, POLD4 was depleted using either of two sets of pooled siRNAs (SI00688695 and SI04189276; Qiagen) as described previously<sup>24</sup>. SMC2 was depleted using either of two previously published siRNAs<sup>19</sup> designated SMC2 siRNA 1 (Cat s20793, Ambion) or SMC2 siRNA 2 (Cat s20794, Ambion). WAPL was depleted using either of two siRNAs designated WAPL siRNA 1 (Cat s22947, Ambion) or WAPL siRNA 2 (Cat s22948, Ambion). SLX4 was depleted using either of two previously published siRNAs<sup>17,18</sup> designated SLX4 siRNA 1 (GAGAAGAACCCUAAUGAAA) or SLX4 siRNA 2 (GCACAAGGGCCCAAGACAA). The Silencer-Select Negative Control siRNA 2 (Cat 4390846, Ambion), the Scrambled shRNA construct<sup>7</sup>, or the Scrambled siRNA control (5'-GACCAAGUUCGUCACUAA-3', Sigma Aldrich), were individually used as negative controls in all transfection experiments, as indicated. Transient transfections using the indicated siRNAs were performed using the Lipofectamine RNAiMax transfection reagent (Life Technologies), according to the manufacturer's instructions.

**Establishment of mammalian expression constructs and stable cell lines.** MUS81 wild type (WT) or MUS81(D338A/D339A) cDNAs<sup>15</sup> were amplified from individual pFastBac vectors using PCR and the following forward (5'-TTCATACCGTCCACCAT-3') and reverse (5'-GTCGACGTAG GCCTTTGAAT-3') primers. PCR products were cloned into pcDNA 3.3 TOPO-TA cloning vectors (Life Technologies). Final constructs were verified by sequencing. The POLD3 siRNA resistant gene was synthesized and cloned into a pcDNA 3.1(+) vector (GeneArt Gene Synthesis, Life Technologies). The MUS81 WT, MUS81(D338A/D339A), or POLD3 WT constructs were transfected into U2OS cells using lipofectamine 2000 (Life Technologies), respectively. Stable cell lines were established and maintained after neomycin (G418) antibiotic selection.

**Cell synchronization and induction of replication stress using low-dose aphidicolin.** Asynchronously growing cells were synchronized in late G2 phase of the cell cycle by incubation with 9 µM RO3306 (ref. 12) (217699; Millipore) for 16 h. Where stated, 0.4 µM aphidicolin (APH; Sigma-Aldrich) (termed low-dose APH) was added for 16 h. Cells synchronized in G2 were subsequently washed with 1 × PBS for 5 min and allowed to progress for a further 30 min into prometaphase before mitotic shake off. For collection of anaphase cells or early G1 daughter cells, pelleted prometaphase cells were re-seeded onto either poly-lysine-coated glass slides (Sigma-Aldrich) or glass coverslips (Sigma-Aldrich), respectively, and incubated at 37°C in an atmosphere of 5% CO<sub>2</sub> and allowed to progress into anaphase for a further 20 min or into the subsequent G1 phase for a further 140 min. Where stated, asynchronously growing cells or cells arrested in G2 were synchronized in prometaphase using 60 ng ml<sup>-1</sup> nocodazole (M1404; Sigma Aldrich) for the indicated times.

**Fixation of cells for immunofluorescence and fluorescence *in situ* hybridization (FISH).** Following the indicated treatments or cell synchronizations, cells for immunofluorescence analysis were fixed and permeabilized for 20 min in PTEMF buffer (20 mM PIPES pH 6.8, 10 mM EGTA, 0.2% Triton X-100, 1 mM MgCl<sub>2</sub> and 4% formaldehyde) at room temperature. Fixed samples were blocked for at least 1 h at room temperature using 3% BSA (Sigma Aldrich) in PBST (containing 0.5% Triton X-100) or stored at 4°C until use. For FISH analysis, metaphase cells or cells in G1 phase were swollen in 75 mM KCl for 5 min and then fixed in methanol:acetic acid (3:1).

## EdU labelling and detection in asynchronous or synchronized cells.

Asynchronously growing cells, treated with or without low-dose APH for 16 h, were pulse labelled using 10 µM EdU (5-ethynyl-2'-deoxyuridine, Life Technologies) for 10, 20 or 30 min followed by simultaneous fixation and permeabilization using PTEMF buffer and subsequent EdU detection using Click chemistry. Cells synchronized in late G2 (+RO3306 ± low-dose APH) were released into early prophase by vigorous washing (3–4 times for up to 5 min) in 1 × PBS (pre-warmed to 37°C). Subsequently cells in early prophase were maintained at 37°C in an atmosphere of 5% CO<sub>2</sub> and released into fresh medium (pre-warmed to 37°C) supplemented with/without 10 µM EdU (Life Technologies) for the indicated time points. This cell synchronization protocol allowed us to detect EdU foci in prometaphase cells following the initiation of EdU incorporation in early prophase. Cells synchronized in prometaphase using nocodazole (60 ng ml<sup>-1</sup>) were pulse labelled with 20 µM EdU and as indicated. The 'Click-chemistry' reaction was performed following the manufacturer's instructions (Click-IT EdU; Alexa fluor 594 or Alexa fluor 488 Imaging Kits, Life Technologies). Immunofluorescence or FISH was carried out after the Click reaction, respectively.

**Immunofluorescence microscopy.** For immunofluorescence analysis, samples fixed on poly-lysine glass slides or glass cover slips were incubated with primary antibodies for up to 2 h followed by 4 washes of 15 min each using PBST (PBS with 0.5% Triton-X 100 (Sigma-Aldrich)) or PBST supplemented with 3% BSA (Sigma-Aldrich). Samples were then incubated with secondary antibodies for 90 min, followed by 4 washes of 15 min using PBST or PBST supplemented with 3% BSA. Air-dried slides or coverslips were mounted using Vectashield mounting medium with DAPI (Vector Laboratories), and were analysed using microscopy.

**Metaphase chromosome spreads, chromosome breakage and chromosome structure analyses.** Cells synchronized in late G2 (using 9 µM RO3306) were washed as stated above and then released into fresh medium (pre-warmed to 37°C) containing 10 µM EdU (final concentration) together with 0.1 µg ml<sup>-1</sup> (final concentration) Karyomax/Colcemid solution in HBBS (Life Technologies) with or without 100 nM (final concentration) of the small molecule inhibitor BI 2536 (PLK1i)<sup>23</sup> for up to 60 min. Metaphase cells were then harvested by mitotic shake-off, pelleted at 600 g for 5 min, and then swollen by incubation in 75 mM KCl (pre-warmed to 37°C) for 5 min at 37°C. Swollen mitotic cells were collected at 600 g for 5 min, fixed using methanol:acetic acid (3:1), and dropped onto pre-hydrated glass slides (Thermo Scientific) and aged for up to 2 days followed by EdU detection, FISH analysis and chromosome structure analysis. Chromosomes were stained using DAPI (Vectashield, Vector Labs).

**Fluorescence *in situ* hybridization (FISH).** DNA used for FISH probes were prepared from the following BAC clones: RP11-149J4 (GenBank: AC098480.2) or RP11-94D19 (GenBank: AC096917.3) for FRA3B, and RP11-264L1 (GenBank: AC046158.6) for FRA16D. BAC DNAs were labelled using the BioNick labelling system (Invitrogen) or DIG-nick translation mix (Sigma Aldrich). FISH was carried out using standard procedures. Briefly, samples were treated with RNase A (0.1 mg ml<sup>-1</sup>), followed by dehydration in ethanol series (50%, 75%, 90%) for 2 min each. Samples were then denatured in 70% formamide at 75°C for 3 min, and dehydrated again in ice-cold ethanol series (50%, 75%, 90%) for 2 min each. FISH-probe hybridization was carried out at 37°C for up to 72 h. Biotin-conjugated probes were detected using avidin-FITC (Life Technologies) and biotin-labelled anti-avidin (Vector Labs). DIG-conjugated FISH probes were detected using antidigoxigenin-rhodamine (Roche). Whole chromosome FISH was carried out using FISH paint probes corresponding to chromosomes 3 or 16 using directly labelled green fluorescent FISH probes purchased from Cytocell (Cat LPP03G and LPP16G, respectively) according to the manufacturer's instructions. Telomere FISH was carried out using the Telomere PNA FISH Kit/FITC (Cat #K5325; DAKO) according to the manufacturer's instructions. Slides were mounted using Vectashield mounting medium with DAPI, and visualized by microscopy.

**Microscopy analysis and live cell imaging.** Images were captured and quantified using a Nikon Eclipse 80i microscope (Nikon, ACT-1, v2.62), Olympus BX63 microscope, or an LSM 700 confocal microscope (Zeiss) equipped with 405 nm, 488 nm and 555 nm lasers, Zeiss imaging software (v7, 2011) and ImageJ analysis software. Live cell images were captured every minute for 4 h using either 63 × (1.2 NA) or 100 × (1.46 NA) objectives.

**Protein quantification and western blot analysis.** Whole cells were lysed using cell extraction buffer (10 mM Tris-HCl, pH 7.4, 100 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1 mM NaF, 20 mM Na<sub>4</sub>P<sub>2</sub>O<sub>7</sub>, 2 mM Na<sub>3</sub>VO<sub>4</sub>, 1% Triton X-100, 10% glycerol, 0.1% SDS, 0.5% deoxycholate), complemented with PIC (Protease Inhibitor Cocktail, Cat #11836153001, Roche) and PMSF (Phenyl-Methanesulfonyl Fluoride Solution; Sigma Aldrich). For subcellular fractionation, cell pellets were incubated in PTEMF buffer (complemented with PIC) or 1 × RIPA buffer (Sigma Aldrich) for 20 min at 4°C on a rotating wheel at low speed. Nuclei were then harvested at 10,000 g for 10 min at 4°C. The supernatant containing the soluble

(cytosolic + nuclear) fraction was saved and stored at 4°C. Subsequently, the insoluble chromatin pellet was reconstituted using cell extraction buffer (complemented with PIC and PMSF) and incubated for 30 min on ice with frequent vortexing. Re-suspended cells, supernatant comprising the soluble fraction, or extracted insoluble chromatin fraction were sonicated four times using a 10 s on/off pulse in a Soniprep 150 Plus sonicator (MSE) or in a Bioruptor Pico sonicator (Diagenode). The chromatin-bound fraction was resolved briefly at 2,000g using centrifugation at 4°C. After protein quantification, samples were boiled for 10 min in SDS loading buffer (1:1 ratio). An aliquot (up to 40 µg) of the resulting sample was run on a SDS-PAGE gel and then transferred to a Hybond-PVDF membrane (Amersham Pharmacia). The membrane was blocked in 5% milk for 1 h at room temperature and incubated overnight with the appropriate primary antibody re-suspended in 5% milk in PBST at 4°C. Following several washes using PBST, the membrane was incubated for 2 h at room temperature with the appropriate secondary antibody re-suspended in 5% milk in PBST at room temperature, followed by 4 washes using PBST. ECL (Amersham) amplification was used for signal detection. Quantifications of western blots were done using Image J/FIJI or Amersham Imager 600 and software (GE Lifesciences).

**Primary and secondary antibodies for western blotting.** Primary antibodies used were MUS81 (1:500, ab14387, MTA30 2G10/3; Abcam), EME1 (1:500, ab88878, Abcam), FANCD2 (1:1,000, NB100-182; Novus), POLD3 (1:500, H00010714-M01; Abnova), POLD4 (1:250, H00057804-M01A; Abnova), PCNA (1:1,000, sc-56; F-2 Santa Cruz), phospho-CHEK1 (Ser 345) (#2341; Cell Signaling), CHEK1 (1:1,000, NCL-CHEK1; Clone DCS-310.1, Leica Biosystems), phosphorylated-histone H3 (Ser10) (PH3S10, indicates mitotic chromatin, 1:1,000, 06-570; Millipore), anti-phospho-Ser/Thr-Pro MPM-2 (indicates mitotic chromatin, 1:1,000, 05-368; Millipore), SLX4 (1:1,000, S714C raised in sheep against residues 1535–1834, a gift from J. Rouse, University of Dundee, UK), polymerase  $\delta$  catalytic subunit (1:1,000, ab10362, 607, Abcam), MCM2 (ab4461; Abcam), MCM3 (1:1,000, ab4460; Abcam), MCM7 (ab2360; Abcam) and actin (1:1,000, A3853, AC-40, Sigma Aldrich). Secondary antibodies used for western blotting were: anti-mouse IgG (whole molecule peroxidase conjugate, 1:2,000, A4416, Sigma Aldrich), anti-rabbit IgG (whole molecule, F(ab')<sub>2</sub> fragment—peroxidase antibody produced in goat, 1:2,000, A6667, Sigma Aldrich) and anti-sheep IgG (whole molecule—peroxidase antibody produced in donkey, 1:2,000, A3415, Sigma).

**Primary and secondary antibodies for immunofluorescence analysis.** Primary antibodies used were MUS81 (1:100, ab14387, MTA30 2G10/3; Abcam), FANCD2 (1:400, NB100-182; Novus), POLD3 (1:100, H00010714-M01; Abnova), PICH/ERCC6L (1:50, 04-1540; Millipore), 53BP1 (1:200, sc22760; Santa Cruz), HEC1 (1:100, ab3613; 9G3; Abcam), cyclin A (1:200, 6E6, ab16726; Abcam), CENPA (1:100, ab13939, 3-19; Abcam), TRF1 (1:100, ab10579, TRF-78; Abcam), RAD51 (1:100, sc-8349; Santa Cruz), TRF2 (1:100, sc-9143; Santa Cruz) phosphorylated-histone H3 (Ser10) (PH3S10, indicates mitotic chromatin, 1:100, 06-570; Millipore), anti-phospho-Ser/Thr-Pro MPM-2 (indicates mitotic chromatin, 1:100,

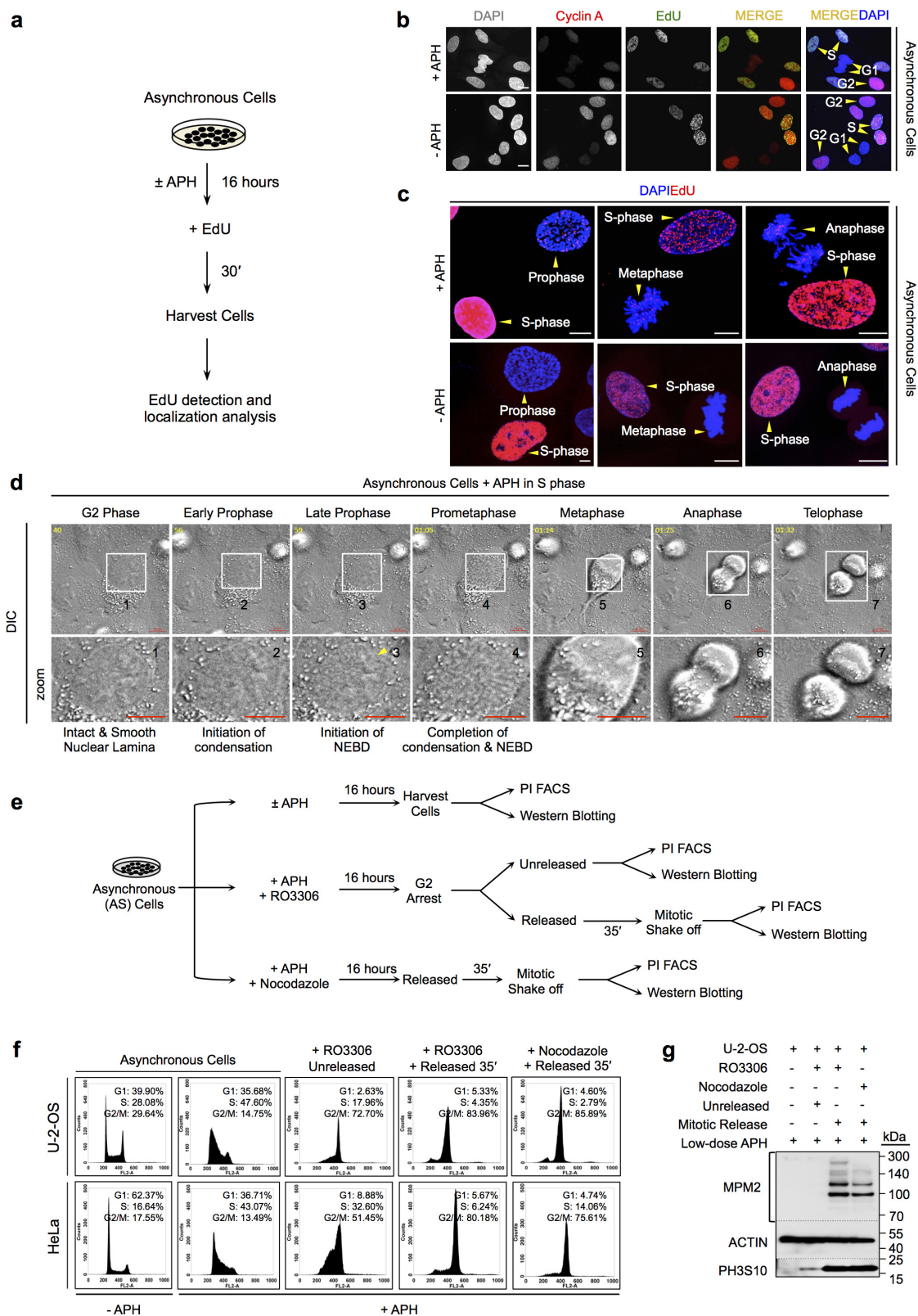
05-368; Millipore) and SLX4 (1:100, S713C, raised in sheep against residues 381-680, a gift from J. Rouse, University of Dundee, UK). Secondary antibodies used for immunofluorescence analysis were the Alexa Fluor 488 goat anti-rabbit IgG (1:500, A-11008; Invitrogen), Alexa Fluor 546 donkey anti-sheep IgG (1:500, A-21098; Invitrogen), Alexa Fluor 488 goat anti-mouse IgG (1:500, A-110291; Invitrogen) and Alexa Fluor 568 donkey anti-mouse IgG (1:500, A-10037; Invitrogen).

**Flow cytometry.** After the indicated treatments or cell synchronization, harvested cells were fixed using ice-cold 70% ethanol (added drop wise) and stored at -20°C for a minimum of 2 h. Fixed cells were rinsed twice in 1 × PBS, and then incubated in 1 × PBS supplemented with 2 mg ml<sup>-1</sup> propidium iodide (Invitrogen) and 10 mg ml<sup>-1</sup> RNase A for 45 min at 37°C. Fluorescence-activated cell sorting (FACS) analysis for each sample was carried out on a FACSCalibur flow cytometer (BD Biosciences). Up to 50,000 cells were analysed for each condition, and data analysis was performed using Cell Quest Pro software.

**Nondisjunction and colony formation assays.** Cells were treated with (+APH) or without (-APH) low-dose APH (0.4 µM) and synchronized at G2 using RO3306 (9 µM) for 16 h. Cells were then released into mitosis in fresh medium with or without high-dose (2 µM) APH for 30 min (incubated at 37°C and 5% CO<sub>2</sub>). Prometaphase cells were then shaken off, pelleted and re-seeded either into six-well plates to determine viability or onto poly-lysine coated glass slides (in the presence of 3 µM cytochalasin B (Sigma Aldrich)). The cells re-seeded onto poly-lysine-coated glass slides were maintained at 37°C in an atmosphere of 5% CO<sub>2</sub> and allowed to progress into the subsequent G1 phase in the presence of cytochalasin B for another 145 min, to generate cytokinesis-blocked, binucleated cells in G1. These cells were processed by methanol:acetic acid (3:1) fixation and subjected to FISH analysis. Cells in six-well plates were maintained at 37°C in an atmosphere of 5% CO<sub>2</sub> and were allowed to form colonies over a period of approximately 10 days. Colonies were stained using crystal violet solution (0.1% Brilliant blue R (Sigma-Aldrich), 7% acetic acid, 50% ethanol) and quantified using ImageJ/FIJI analysis software.

**TUNEL assays.** TUNEL assay was performed using the DeadEnd Fluorometric TUNEL System, (Promega) following the manufacturer's instructions. Briefly, following indicated treatments, cells were fixed in paraformaldehyde (4%) at 4°C for 25 min, followed by washing with PBS. Slides were immersed in 0.2% Triton X-100 in PBS for 5 min and then washed again. Subsequently, samples were incubated in equilibration buffer at room temperature for 5–10 min. For labelling, 50 µl of TdT reaction mix was added to the cells on an area no larger than 5 cm<sup>2</sup>, and the slides were incubated for 1 h at 37°C in a humidified chamber. The reaction was terminated by immersing the slides in 2 × SSC buffer. All nuclei were stained with DAPI. Green fluorescence was detected by fluorescence microscopy.

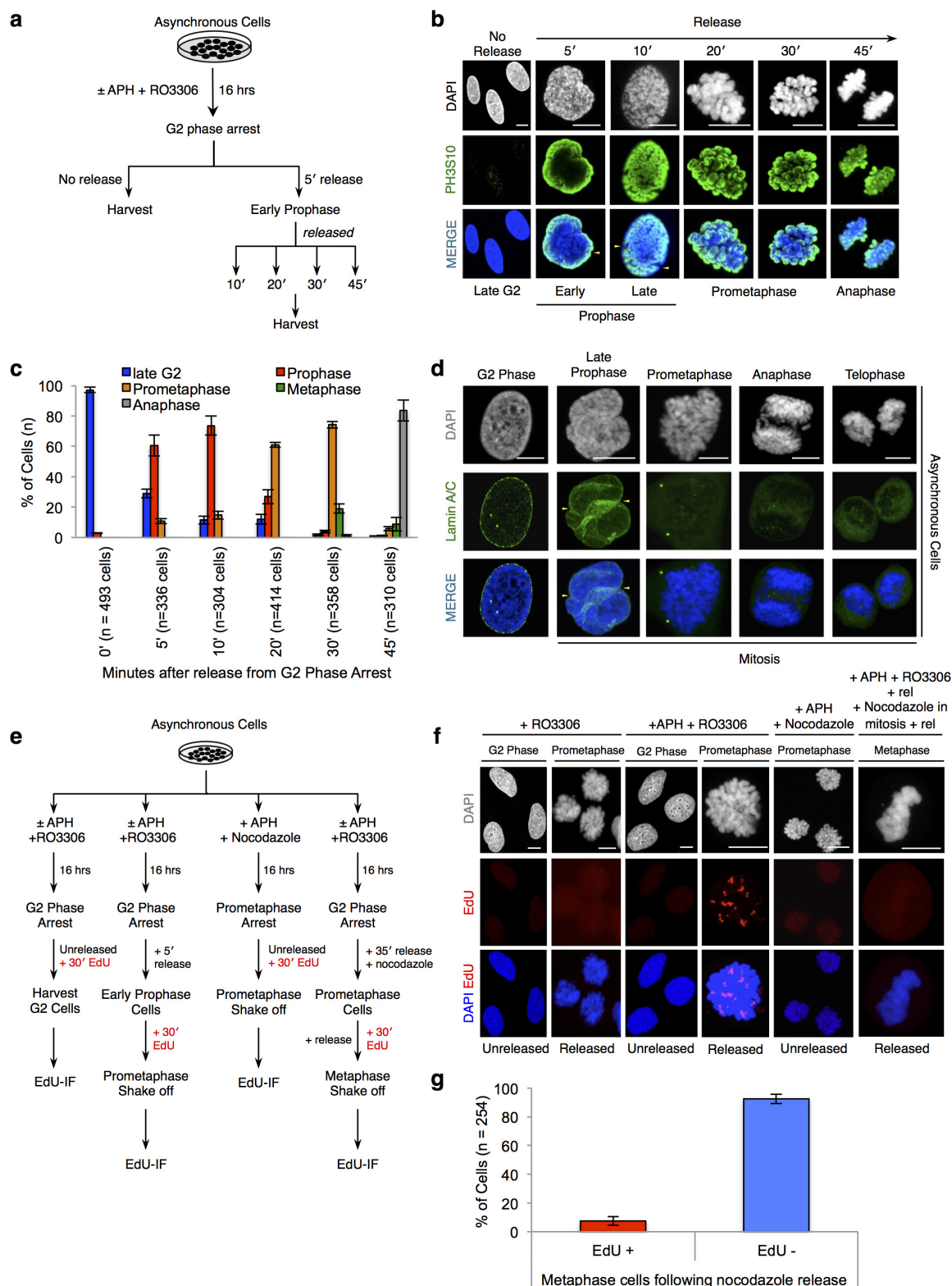
**Statistical analysis.** At least three independent experiments were carried out to generate each data set and statistical significance in each case was calculated using Pearson's  $\chi^2$  test as indicated in each figure legend.



**Extended Data Figure 1 | Replicative stress activates DNA synthesis in mitosis.** **a, b**, Experimental workflow (**a**) and representative images (**b**) showing asynchronously growing U2OS cells pre-treated with low-dose APH (+ APH) or without APH (− APH) followed by EdU labelling for 30 min. Nuclei were stained with DAPI (blue). S-phase cells are indicated as being positive for cyclin A (red) and EdU incorporation (green). Cells in G2 (cyclin A positive) and G1 daughter cells (cyclin A negative) show negligible EdU staining. **c**, Uncropped images from Fig. 1a showing EdU foci (red) formation in condensed mitotic nuclei stained using DAPI (blue). Scale bars (**b, c**), 10  $\mu$ m. **d**, Still images from Supplementary Video 1 showing

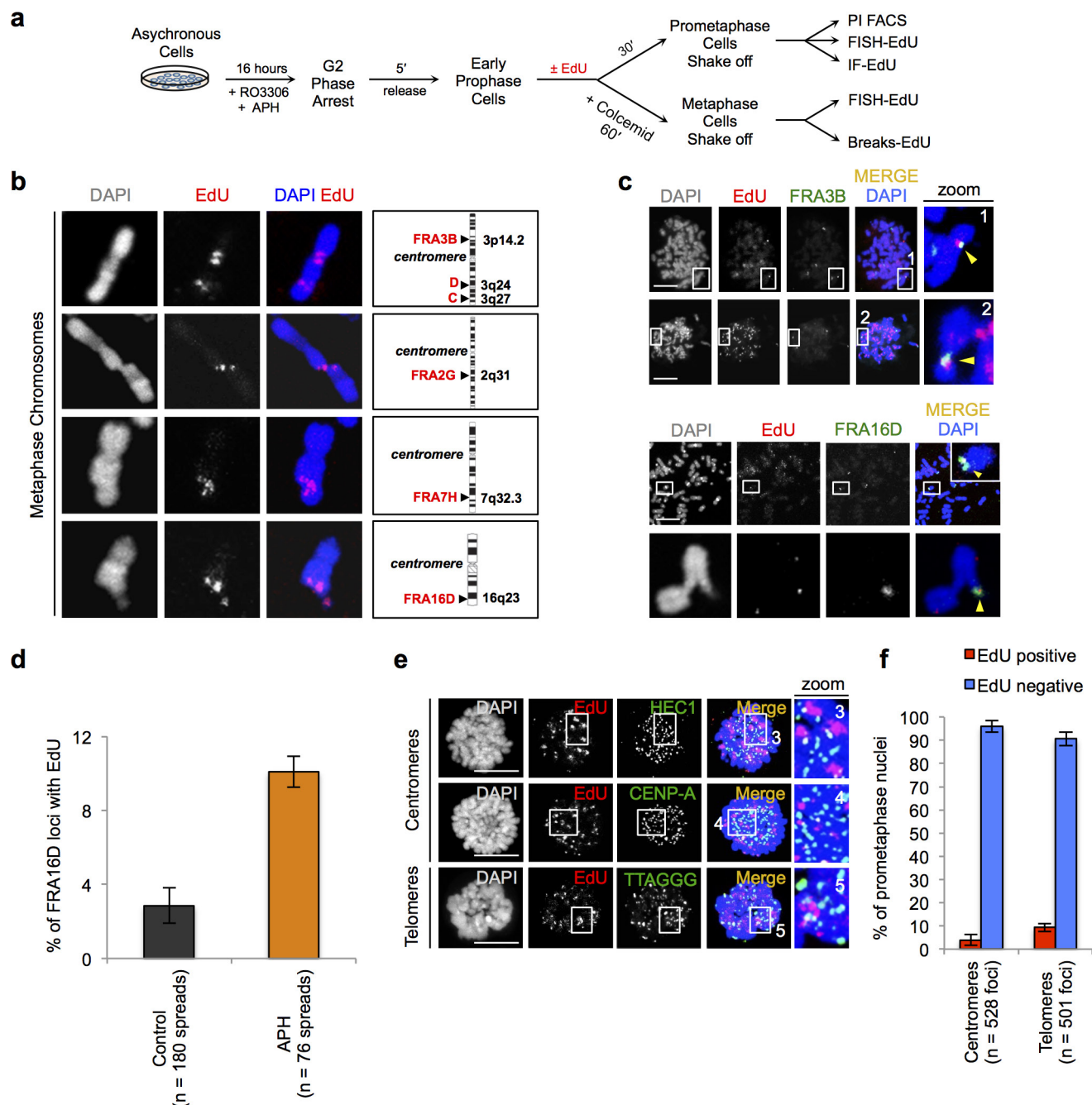
representative individual frames from a live cell imaging experiment ( $n = 3$  live cell imaging experiments) at various time points. The yellow arrow (frame 3, from the left) shows the initiation of nuclear envelope breakdown (NEBD) and characteristic invaginations in the nuclear lamina in a late prophase cell with partially condensed chromosomes. Selected areas (white boxes) are enlarged below. Scale bars, 10  $\mu$ m. **e**, Experimental workflows used to follow cells through mitosis and into the next cell cycle and as indicated. **f**, FACS analysis of U2OS and HeLa cells as indicated above. **g**, Western blot analysis for the mitosis-specific MPM2 epitope-containing antigens (top panel) and PH3S10 (bottom panel). Actin was used as a loading control.





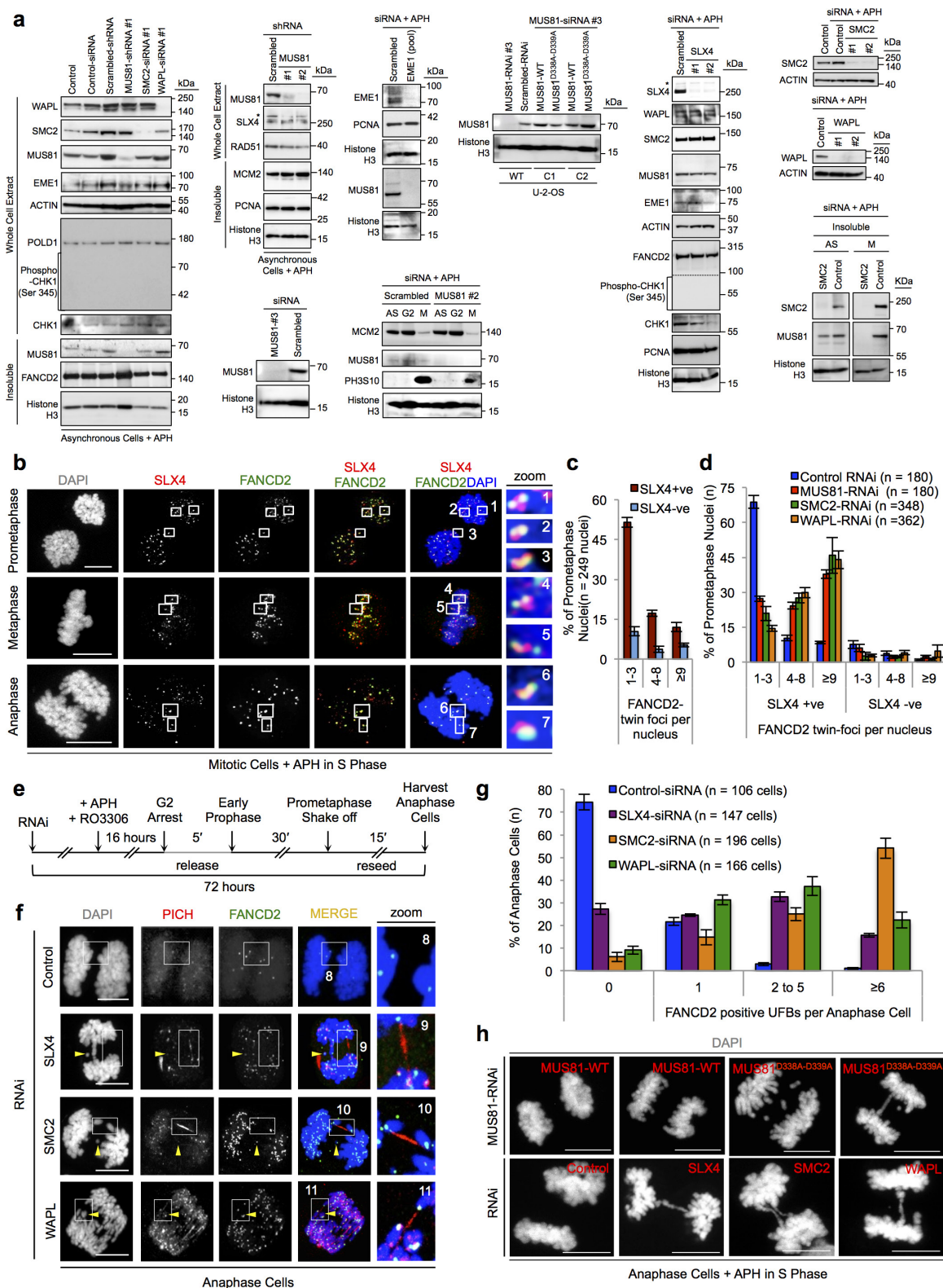
**Extended Data Figure 2 | Assessment of mitotic progression and EdU incorporation following a reversible arrest of cells in very late G2 using the RO3306 CDK1 small molecule inhibitor. a–c,** Experimental workflow used (a) and representative images (b) and quantification (c) showing the progression of cells through mitosis using PH3S10 (green) as a marker of entry into mitosis and chromosome condensation following a reversible G2 arrest. Nuclei and chromosomes are stained using DAPI (blue). **d,** Representative images showing the progressive disappearance of the nuclear lamina (lamin A/C, green) as cells progress into mitosis while condensing their nuclear genome. Yellow arrows indicate invaginations in the nuclear lamina upon entry into mitosis. Un/condensed nuclei were

stained using DAPI (blue). **e, f,** Experimental workflows used (e) and representative images (f) showing DNA synthesis (EdU foci, red) in late G2 phase arrested cells or in cells released from this arrest into prometaphase or metaphase and as indicated. Low-dose APH (+APH) and/or EdU were added to cells where indicated. Intact or partially condensed nuclei are stained using DAPI (blue). **g,** Quantification of EdU positive (EdU<sup>+</sup>) or EdU negative (EdU<sup>-</sup>) metaphase cells following a release from the nocodazole arrest (e, right column). Scale bars (b, d, f), 10  $\mu$ m. *n* = number of cells quantified in each case as indicated. Data are means of at least three independent experiments. Error bars represent s.e.m.



**Extended Data Figure 3 | Replicative stress activates DNA synthesis at common fragile sites (CFSs) in mitosis.** **a, b**, Experimental workflow used (**a**) and representative images (FRA3B cropped from Fig 1d) (**b**) showing examples of chromosome breaks/gaps harbouring known CFS loci (ideograms, black arrows) and that are positive for EdU (red). Chromosomes were stained using DAPI (blue). Images were captured using a 60 $\times$  objective. **c**, Representative images showing EdU FISH analysis to assess the co-localization (yellow arrowheads) between EdU foci (red) and biotin-labelled FRA3B or FRA16D FISH signals (green) in

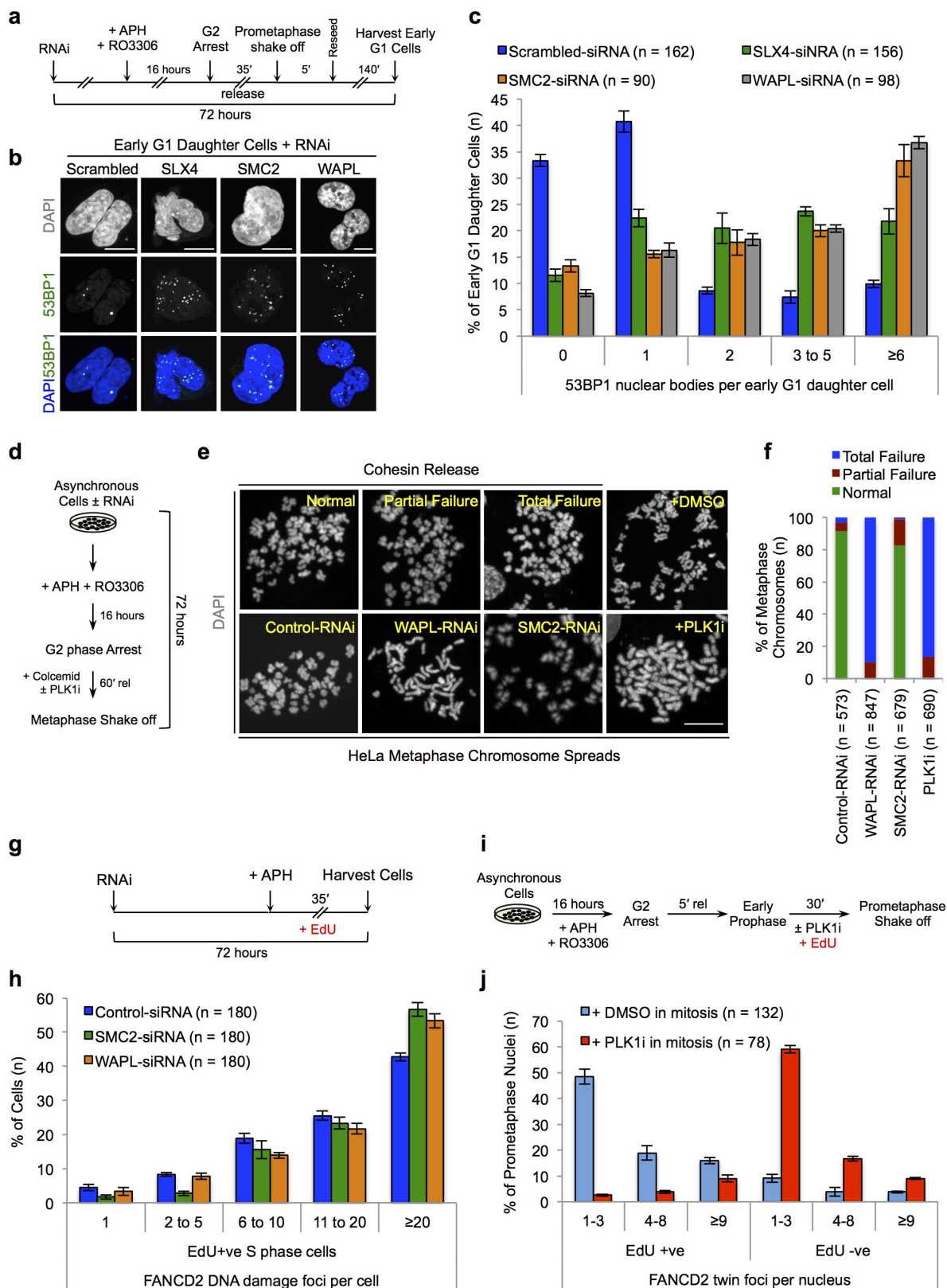
individual metaphase spreads. Chromosomes were stained using DAPI (blue). Selected regions of interest are magnified and shown in numbered panels. **d**, Quantification of co-localized EdU and FRA16D foci from panel **c** in cells treated with (APH) or without (control) low-dose APH. **e, f**, Representative images (**e**) and quantification (**f**) of EdU foci (red) at centromeres or telomeres (green) in DAPI-stained mitotic nuclei (blue). Selected regions of interest are magnified and shown in numbered panels below. Scale bars (**c, e**), 10  $\mu$ m. Data are means of at least three independent experiments. Error bars represent s.e.m.



**Extended Data Figure 4 | Catalytically active MUS81, SLX4, WAPL and SMC2 are collectively required for sister chromatid disjunction in anaphase.** **a**, Representative western blots of the indicated proteins (left) after siRNA treatments (shown above). Actin and histone H3 were used as whole-cell lysate and nuclear loading controls, respectively, and PH3S10 indicates mitotic chromatin. **b**, **c**, Representative images (b) and quantification (c) of co-localized SLX4 foci (red) and FANCD2 twin foci (green) in mitotic cells (DAPI, blue) after pre-treatment of cells with low-dose APH (+APH). Selected regions of interest are magnified and shown in numbered panels. **d**, Same as panel c, except following the indicated RNA interference (RNAi). **e**–**g**, Experimental workflow used (e), representative

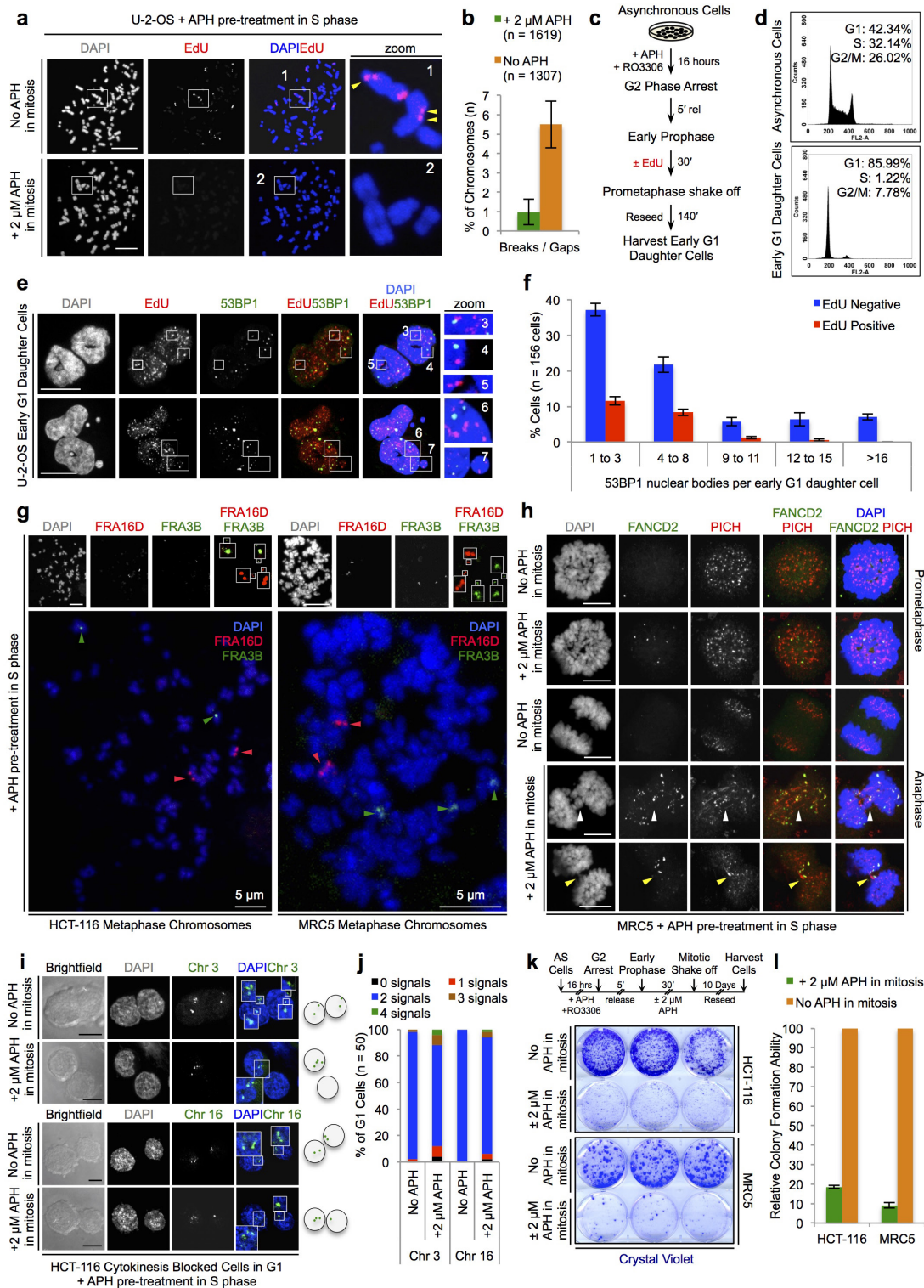
images (f) and quantification (g) of anaphase cells stained for PICH (red) and FANCD2 (green) after depletion of the proteins indicated on the left. Selected regions of interest are magnified and shown in numbered panels. Yellow arrows show lagging chromatids. DNA was stained with DAPI (blue). **h**, Representative images showing DAPI-positive bulky chromatin bridges in anaphase cells after depletion of indicated proteins and pre-treatment of cells with low-dose APH. In the upper panel, MUS81 RNAi-treated cells were reconstituted with either wild-type MUS81 (MUS81 WT) or a nuclease-dead MUS81 (MUS81(D338A/D339A)) version, as indicated in red. Scale bars (b, f, h), 10  $\mu$ m. Data are means of at least three independent experiments. Error bars represent s.e.m.





**Extended Data Figure 5 | Early mitotic events promote DNA synthesis at CFSs in early mitosis and reduce the transmission of errors to daughter cells.** **a–c**, Experimental workflow used (**a**), representative images (**b**) and quantification (**c**) of 53BP1 nuclear bodies (green) in G1 daughter cells (DAPI, blue) after the indicated siRNA treatments (shown above). **d–f**, Experimental workflow used (**d**), representative metaphase chromosome spreads (stained using DAPI) (**e**) and quantification (**f**) of sister chromatid cohesion release after depletion of the proteins indicated

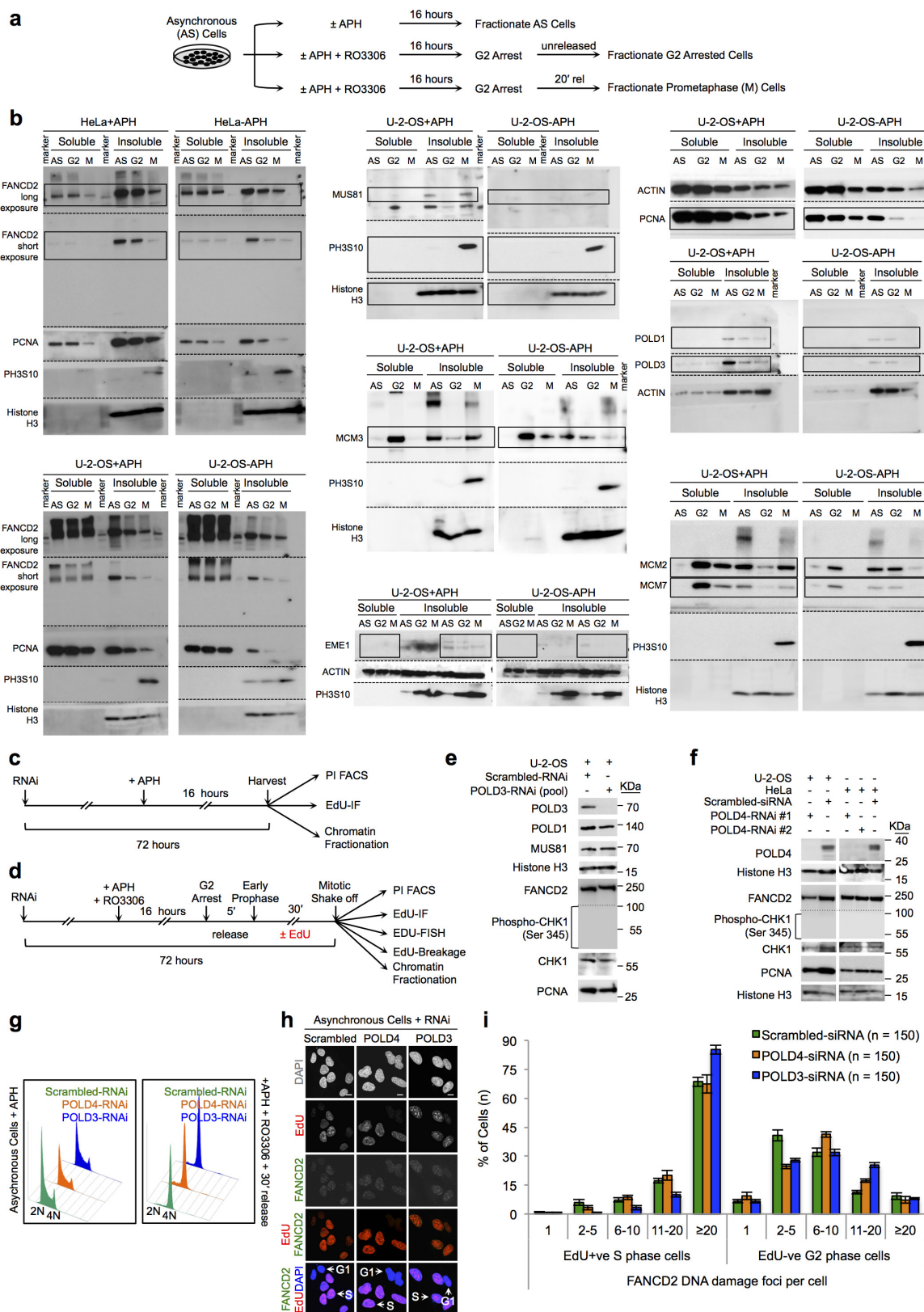
or following inhibition of PLK1 using the small molecule inhibitor BI-2536 (PLK1i) in mitosis. **g, h**, Experimental workflow used (**g**) and quantification (**h**) of FANCD2 DNA damage foci in EdU<sup>+</sup> S phase cells after RNAi and APH pre-treatment (+APH) as indicated. **i, j**, Experimental workflow used (**i**) and quantification (**j**) of EdU-positive (EdU<sup>+</sup>) or EdU-negative (EdU<sup>-</sup>) FANCD2 twin foci in prometaphase cells after exposure to PLK1i in mitosis. Scale bars (**b, e**), 10  $\mu$ m. Data are means of at least three independent experiments. Error bars represent s.e.m.



### Extended Data Figure 6 | Administration of high-dose APh in mitosis abolishes DNA synthesis in mitosis and decreases cell survival.

**a, b**, Representative images (**a**) and quantification (**b**) of EdU incorporation at breaks/gaps on metaphase chromosomes (DAPI, blue) after low-dose aphidicolin pre-treatment in S phase. **c–f**, Experimental workflow (**c**), representative FACS analysis profiles (**d**), representative EdU immunofluorescence images (**e**), and quantification (**f**) of a lack of co-localization between sites of EdU incorporation (labelled in the previous mitosis; red) and 53BP1 nuclear bodies (green) in G1 daughter cells. Selected regions of interest are magnified and shown in numbered panels. **g**, Representative FISH images showing localization of FRA16D (red signals) and FRA3B (green signals) on metaphase chromosomes (DAPI, blue). **h**, Representative images showing CFS-associated FANCD2 twin foci

(green) and PICH staining (red) in prometaphase and anaphase MRC5 cells treated with or without high-dose APh ( $\pm 2 \mu\text{M}$  APh) in mitosis and as indicated. White and yellow arrows indicate lagging chromosomes and anaphase bridges, respectively. **i, j**, Representative images (**i**) and quantification (**j**) showing FISH analysis using directly labelled 'paint' probes for chromosome 3 (green) or chromosome 16 (green) in cytochalasin B arrested binucleated cells (DAPI; blue) treated or not with low-dose aphidicolin (+ APh) in S phase. **k, l**, Experimental workflow used (**k**, top), representative crystal violet stained colonies (**k**, bottom) and quantification (**l**) of colony formation assay used to assess cell viability as indicated in HCT-116 or MRC5 cell lines. Scale bars,  $10 \mu\text{m}$  unless indicated otherwise. Data are means of at least three independent experiments. Error bars represent s.e.m.

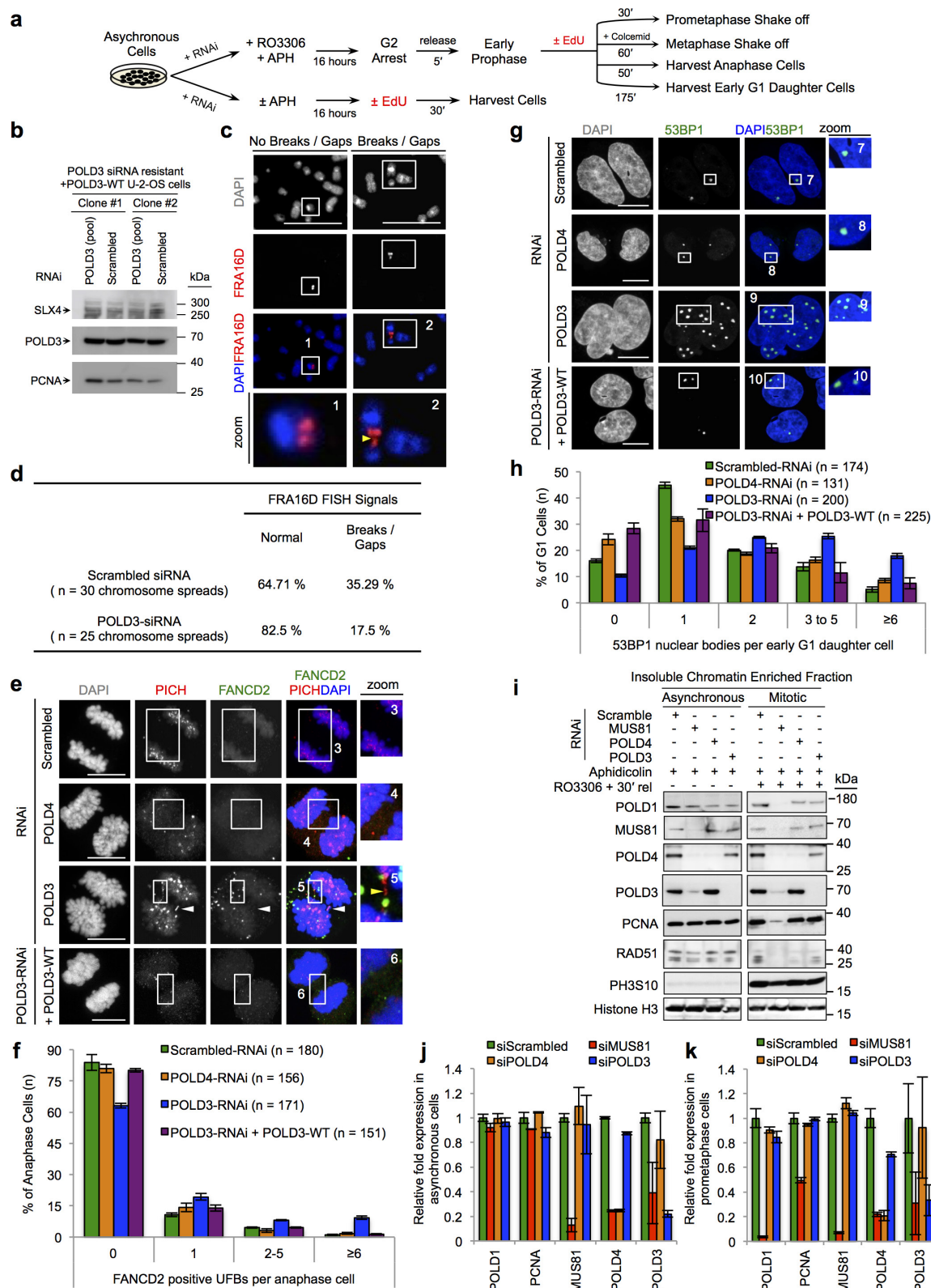


### Extended Data Figure 7 | DNA repair proteins recruited to regions synthesizing DNA in early mitosis after replicative stress.

**a, b**, Experimental workflows (a) and representative extended western blots of the indicated proteins (left, b) and in the indicated cell lines (top) from Fig. 4a. Actin and histone H3 were used as loading controls for soluble (cytosolic and nuclear) and insoluble (chromatin bound) fractions, respectively. PH3S10 indicates mitotic chromatin fractions. Three cell populations are shown: asynchronous (AS), G2-arrested (G2)

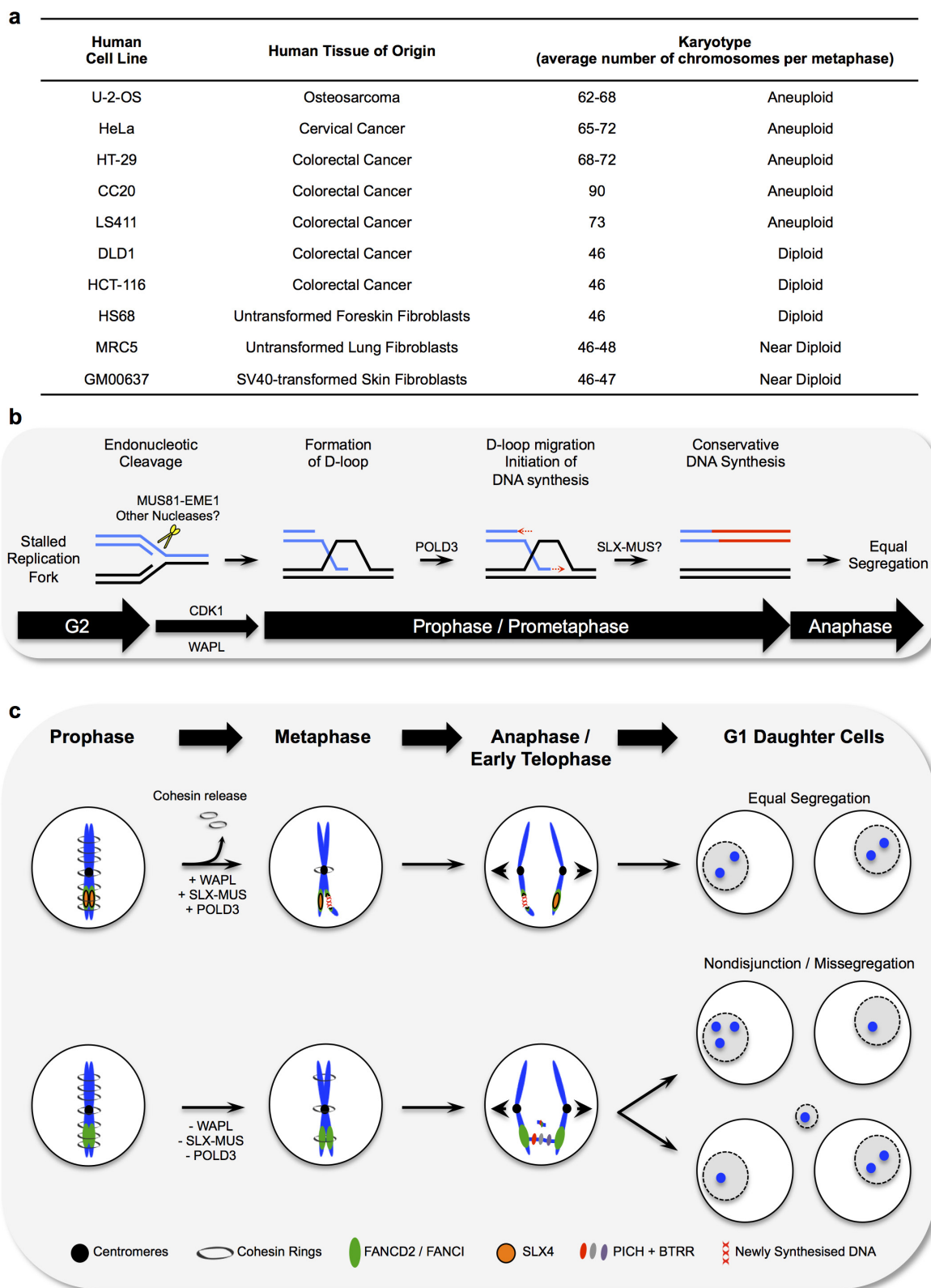
and prometaphase (M). **c-f**, Experimental workflows used (c, d) and western blots (e, f) showing POLD3 or POLD4 levels after either POLD3 (e) or POLD4 (f) siRNA treatments (shown above) in the indicated cell lines. Histone H3 was used as a nuclear loading control. **g**, FACS traces of asynchronous (left) and prometaphase (right) cells following RNAi as indicated. **h, i**, Representative images (h) and quantification (i) of FANCD2 DNA damage foci in EdU<sup>+</sup> S-phase cells or EdU<sup>-</sup> G2 phase cells after RNAi and APH treatment (+APH).





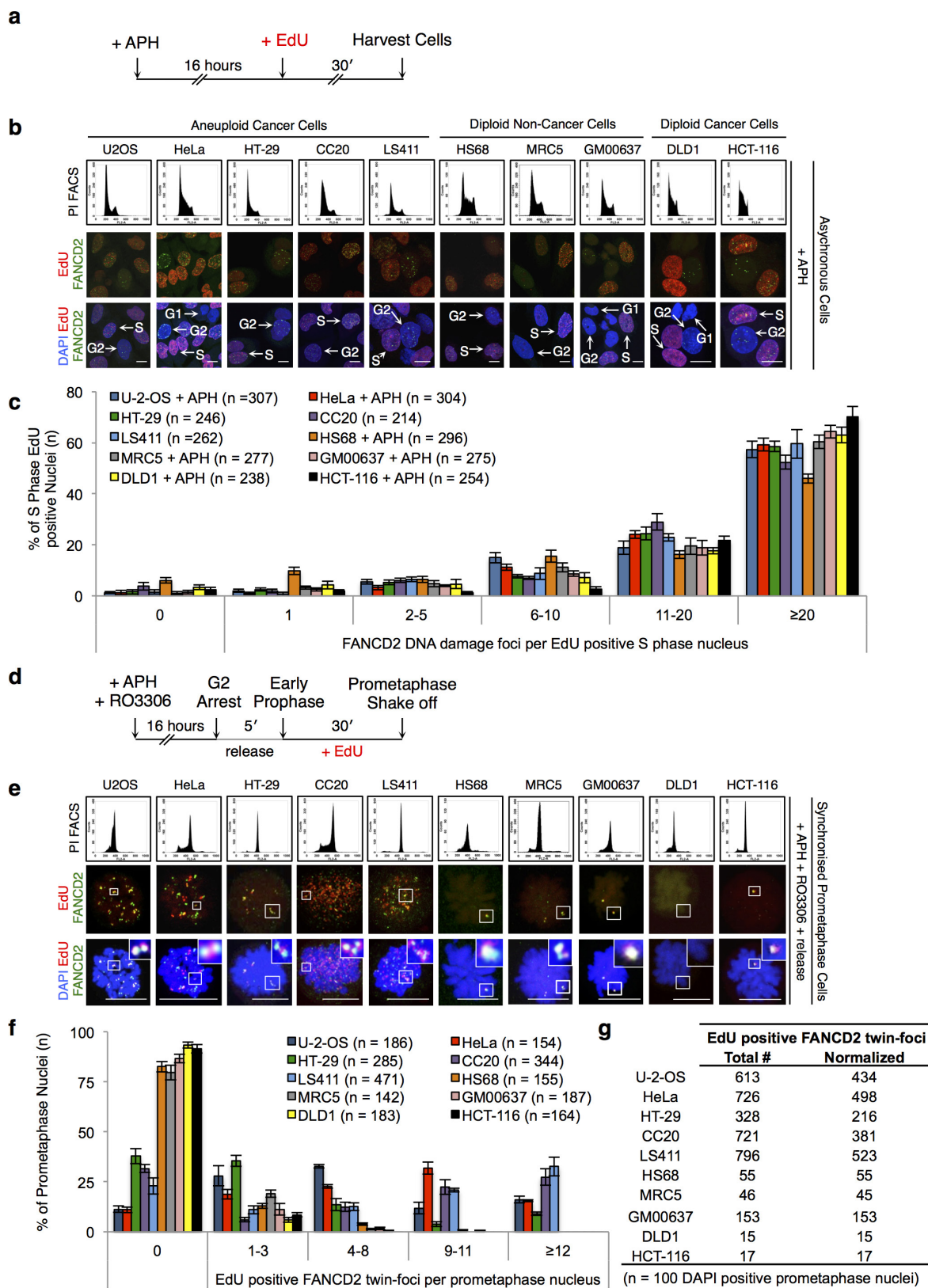
**Extended Data Figure 8 | POLD3 promotes faithful chromosome segregation and maintains genome stability.** **a**, Experimental workflow used. **b**, Representative western blot analysis of the proteins as indicated (left) from the whole-cell lysates obtained from two POLD3-siRNA-resistant wild-type POLD3 U2OS stable clones following the indicated RNAi treatments. **c**, **d**, Representative FISH images of isolated metaphase chromosomes (**c**) and quantification (**d**) of chromosome breaks/gaps at the FRA16D CFS locus (red). **e**, **f**, Representative images (**e**) and quantification (**f**) of UFOs stained using PICH (red) and FANCD2 foci at their bridge termini (green) in anaphase cells (stained using DAPI) (blue) and following

the siRNA treatments as indicated (left). Selected regions are magnified and shown in separate numbered panels. **g**, **h**, Representative images (**g**) and quantification (**h**) of 53BP1 nuclear bodies (green) in early G1 daughter cells (stained using DAPI) (blue) following the siRNA treatments as indicated (left). **i**–**k**, Representative western blots (**i**) and their quantification (**j**, **k**) of the indicated proteins (left) in chromatin-bound nuclear fractions from asynchronous or prometaphase (mitotic) cells following the RNAi treatment as indicated above. Histone H3 was used as a chromatin loading control and PH3S10 indicates mitotic chromatin. All scale bars, 10  $\mu$ m. Data are means of at least three independent experiments. Error bars represent s.e.m.



**Extended Data Figure 9 | Cell lines analysed, and proposed models summarizing and extending the main findings of this study.** **a**, A list of cultured human cell lines analysed in this study. **b**, DNA-level model showing how a stalled replication fork at a CFS might be repaired after replication stress. This would first require cleavage of the stalled replication fork, which we propose utilizes the MUS81–EME1 endonuclease and/or other nucleases. The subsequent formation of a D-loop structure would initiate POLD3-dependent conservative DNA synthesis, akin to BIR, in early mitosis. **c**, A model at the chromosome level to indicate how MUS81- and POLD3-dependent DNA synthesis is triggered at CFS

(marked by SLX4 and FANCD2 twin foci), and how it is dependent on the collective action of chromosome condensation and WAPL-mediated cohesin release in prophase (diagrams in top panel). DNA synthesized in this pathway manifests as uncondensed regions that nevertheless facilitate chromosome disjunction (diagrams in top panel). In contrast, if WAPL-dependent cohesion dissolution in prophase or chromosome condensation is defective, or when MUS81 or POLD3 function is compromised, a reduction in CFS expression is observed, associated with an increase in chromosome non-disjunction/mis-segregation events, which can activate cell death and increase the frequency of aneuploidy (diagrams in bottom panel).



**Extended Data Figure 10 | Aneuploid/CIN<sup>+</sup> cancer cells show elevated levels of DNA synthesis at CFSs in condensed mitotic nuclei.**

**a**, Experimental workflow used to assess the DNA damage response in human cells after replicative stress (+APH). **b**, **c**, Representative images of asynchronous human cells (**b**) and quantification (**c**) showing FACS (top panel) and immunofluorescence analysis for FANCD2 DNA damage foci (green) in EdU-positive (red) S-phase nuclei (bottom two panels). **d**, Experimental workflow used to assess mitotic DNA synthesis at CFS-associated FANCD2 twin foci in human cells after replicative stress

(+APH). **e**, **f**, Representative images (**e**) and quantification (**f**) showing FACS (top panel) and immunofluorescence analysis for EdU foci (red) and CFS-associated FANCD2 twin foci (green) (bottom two panels) in cells arrested in prometaphase. **g**, Total number and normalized values of co-localized EdU foci and FANCD2 twin foci. In each cell line studied, values per 100 cells scored were normalized to a diploid chromosome number. Scale bars (**b**, **e**), 10  $\mu$ m. Data are means of at least three independent experiments. Error bars represent s.e.m.

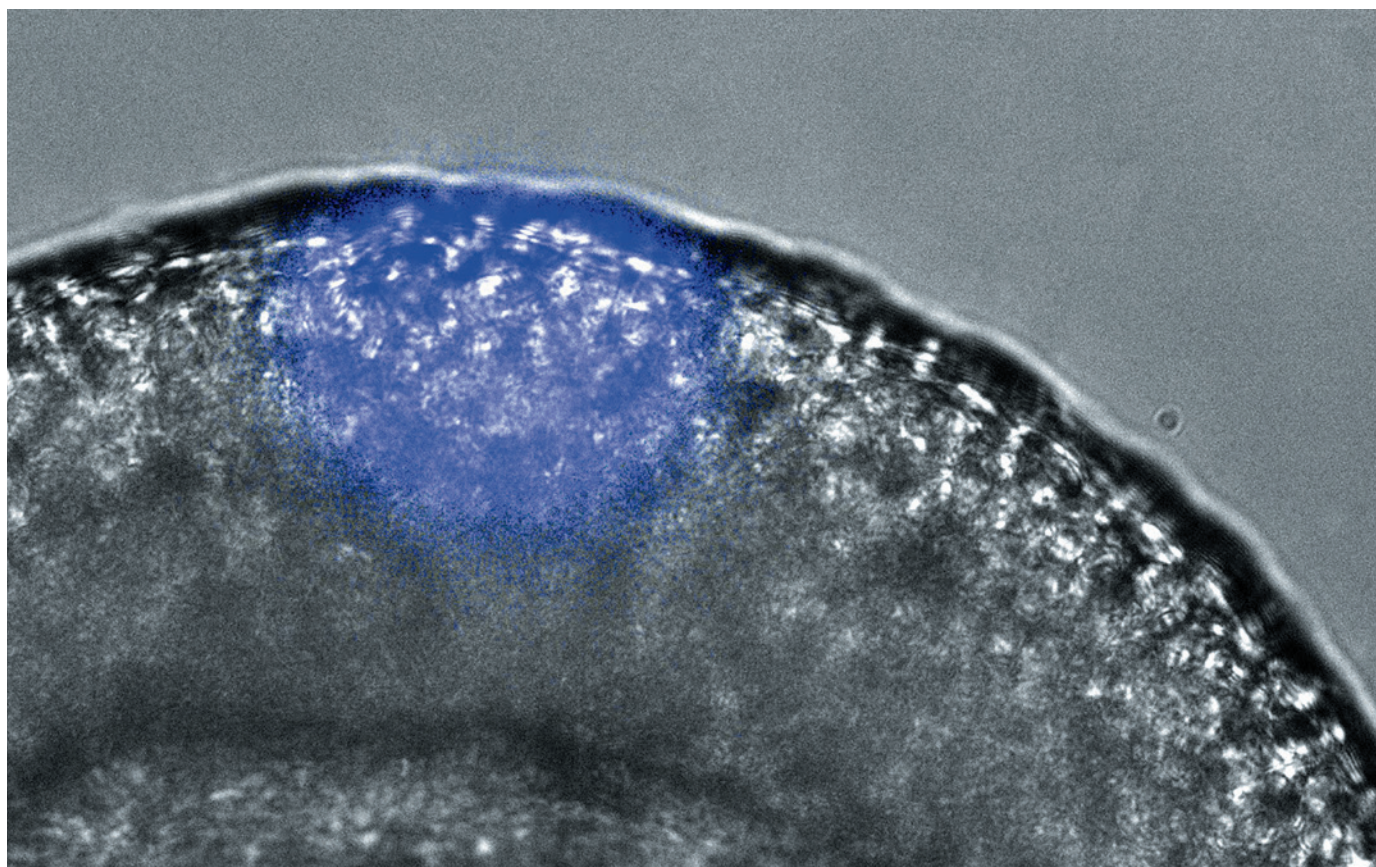


## TECHNOLOGY FEATURE

# MICROMANAGEMENT WITH LIGHT

*The optogenetics techniques that have long been used in neuroscience are now giving biologists the power to probe cellular structures with unprecedented precision.*

ANNA READE, STEPHANIE WOO, ORION WEINER, UCSF



Shining blue light at just one part of an engineered zebrafish embryo enables scientists to selectively activate a light-sensitive transcription factor.

BY AMBER DANCE

Kevin Gardner opens up a mini-fridge-sized incubator and stares at the flashing blue lights inside, a scene he always finds reminiscent of a 1970s New York disco. “There are interesting things happening,” he notes — but rather than the disco lights, he’s talking about events at the microscopic scale.

Gardner is a structural biologist at the City University of New York’s Advanced Science Research Center, where he is a leader in the use of light to control the activity of proteins, a technique known as optogenetics. Thanks to the tools that he and other protein engineers

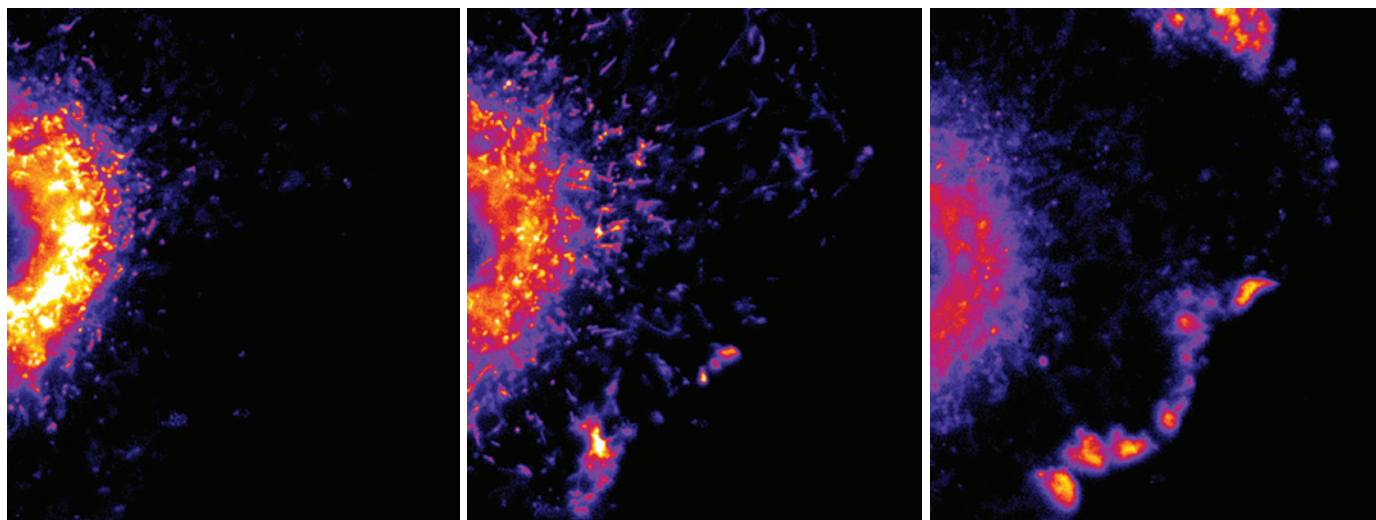
have developed, scientists can now micromanage processes such as cell signalling or movement with an LED or laser flash, rather than just observing them. They can flip proteins on and off, for example, or move organelles back and forth across a cell.

Over the past several years, protein engineers have developed nearly a dozen light-sensitive tools that they can use to accomplish such feats. Some are artificial proteins designed by scientists, but many incorporate modified versions of natural light-sensing proteins. A simple example is the light-oxygen-voltage-sensing (LOV) domain. Found in plants, fungi and some bacteria, it contains a portion that

winds up into a helix. In the dark, this coil tucks in close to the rest of the protein. But under blue light, the helix lets go and lays bare the structures hidden beneath it. Plants and algae use LOV sensors to cover enzymes or DNA-binding proteins, enabling them to regulate activities such as growth towards light or rearrangements of chloroplasts. But scientists can make a custom light-activated protein by choosing what is hidden underneath the LOV coil — the active site of an enzyme, for example.

Light offers important advantages over standard methods of manipulating cellular activity. One such advantage is speed. Chemicals take minutes to enter a cell, whereas light takes ►





LUKAS KAPITEIN, UTRECHT UNIV.

Under blue light, peroxisomes near the nucleus (left) attach to the motor protein kinesin, which drags them to the cell periphery (middle and right)<sup>1</sup>.

► fractions of a second. Thus, cell biologists can probe cellular processes such as signalling pathways or protein movement that take place on time scales of seconds to minutes, says Klaus Hahn, a cell biologist and protein engineer at the University of North Carolina School of Medicine at Chapel Hill. Likewise, a cell or organism with a knocked-out or knocked-down gene gets days, weeks or even longer to adapt to the change, and perhaps activate back-up systems. But if the protein is deactivated by a light switch on the microscope stage, there is no time to compensate — and researchers may see effects that they would not observe with conventional methods. “The cell doesn’t know what hit it,” Hahn says. And the effects can be reversed by simply turning off the light.

Another advantage is that optogenetics offers precise spatial control: instead of flooding every cell in a Petri dish with the same small-molecule treatment, cell biologists can use tightly focused light to flip the switch in just one cell, or even part of a single cell.

Optogenetics flourished initially in neuroscience: light-controlled channels were used to make neurons fire at will. But cell biologists have now embraced the technique enthusiastically. “You’re going to see a ton of papers coming out, in every organism you can think of, using these tools, within the next 12 months,” predicts Jared Toettcher, a bioengineer at Princeton University in New Jersey.

### PROTEIN PARTNERS

One of the most common tricks in optogenetics is to design two proteins that will bind to each other in the presence of light, forming a ‘dimer’. Scientists have been triggering dimer formation with chemicals for some time, but doing it with light is relatively new. The importance of protein–protein interactions in biology makes light-induced dimerization a game-changer, says Chandra Tucker, a biochemist at the University of Colorado School of Medicine in Denver. “If you are very creative,” she says, “you can control

[protein] activities in many different ways.” For example, scientists can tether one of the proteins on a cellular membrane, and leave the other free-floating. When they turn on the light, the mobile partner will be captured by the membrane-bound partner, thus targeting it to that location. Or they can split a single protein into two inactive fragments and reattach them with a light switch to make the functioning version.

Lukas Kapitein, a biophysicist at Utrecht University in the Netherlands, used light-induced dimerization to move individual classes of organelles around like furniture in a house<sup>1</sup>. Scientists have realized lately that cells rely on a certain *feng shui*. For example, when there are plenty of nutrients around, lysosomes — metabolic organelles — hang out near the cell’s edges, promoting the production of new proteins. But when cells are starved, lysosomes retreat to the cell’s interior, where they encourage the cell to start digesting itself<sup>2</sup>.

Organelle location can even affect a cell’s shape. Neurons send out projections called axons to transmit impulses to other neurons, and axons tend to branch into two at spots where mitochondria, the cell’s energy facility, have settled<sup>3,4</sup>.

However, the effects of cellular layout can be difficult to unravel. In the past, cell biologists generally had to rely on wholesale techniques such as dissolving the cytoskeleton or changing the levels of molecular motors that deliver organelles to the right spot — relatively crude processes that tended to move all of the organelles simultaneously. Conversely, Kapitein’s optogenetic method offers the ability to fine-tune the positioning of a single kind of organelle, and it is reversible. The main optogenetic tools he uses are tunable light-inducible dimerization tags (TULIPS), which are based on the LOV photosensor from oats, and an engineered

protein–protein interaction domain based on the common PDZ sequence. The LOV helix hides a small peptide, which, when exposed by blue light, binds to the PDZ domain<sup>5</sup>.

Kapitein started by attaching the LOV domain to three different kinds of organelles: mitochondria, peroxisomes (metabolic sacs) and recycling endosomes that return internalized membrane components to the plasma membrane. Then he hooked the PDZ domain to one of two different kinds of intracellular motors: kinesins, which drag their cargoes to the cell’s perimeter, and dyneins, which tote cargo towards its centre. With a flash of blue light, Kapitein could shuffle specific organelles inward or outward (see ‘Light switch’).

The researchers applied their TULIP set-up to test how endosome location affects axon growth in neurons. They removed the endosomes from the axon tips, which stopped the axons from extending. They shoved in extra endosomes, and the axons grew faster. Thus, as with mitochondria, the position of these organelles affects the cell’s shape.

The same system should work for many kinds of organelle, says Kapitein, allowing scientists to ask previously unanswerable questions about cell layout. He has received dozens of requests for his constructs from cell biologists who want to rearrange their own favourite cell structures. Looking ahead, he wants to find a way to move a single organelle (as opposed to all of the organelles of a particular class) and park it at a desired location.

### SIGNAL OF INTENT

Biologists do not have to reposition entire organelles to make waves in a cell; moving a single protein will do. Many signalling pathways start with the binding of some external factor to receptors on the cell membrane, followed by a cascade of interactions that transfers the information inward from one protein to the next. The end result is some appreciable change in the interior, such as a shift in gene expression. Scientists can

**“With a flash of light, Kapitein could shuffle specific organelles inward or outward.”**

often mimic these effects by identifying proteins involved in the early stages of the pathway and moving them to the plasma membrane. After they hit the membrane, the proteins act as though they have received the external signal and kick off the downstream cascade.

For example, Toettcher and his colleagues used a light-controlled system to study the effects of Ras, a signalling protein that is involved in diverse processes such as cell proliferation and determining cell fate in a developing embryo. This one signalling pathway can mediate such different processes because Ras has a different effect according to when and where in the cell it gets activated — but researchers were unable to investigate this in great detail until they had the optogenetic tools to turn Ras on and off.

Toettcher used the phytochrome B (PhyB)–PIF dimerization system, which optogenetics scientists have borrowed from the plant geneticist's favourite weed, *Arabidopsis*<sup>6</sup>. In the plant, visible red light causes PhyB to bind and activate the PIF transcription factor — a mechanism that *Arabidopsis* uses to turn on genes involved in processes such as germinating seeds or growing away from shade. But unlike other optogenetics systems that simply switch off in the dark, PhyB and PIF stay bound until they are hit with longer-wavelength infrared light. Toettcher hooked PhyB to the plasma membrane, and part of PIF to a Ras activator. When he turned on the red light, Ras would turn on too.

Because he could turn Ras back off with infrared rays, Toettcher could precisely control the timing of its activation over minutes or hours, and this made a difference to what happened downstream. For example, turning on Ras in one cell causes its neighbours to phosphorylate STAT3: a transcription factor that works in various processes such as cell growth and death. Two hours of continuous red light stimulated STAT3 phosphorylation. But 1 hour of red light, 15 minutes of infrared light and another hour of red light did not, Toettcher says. Although Ras activation totalled two hours in both cases, the cell could

tell the difference, and responded accordingly<sup>7</sup>. The researchers do not know precisely what use STAT3 is being put to after extended Ras signalling, but they surmise that this kind of system would allow a cell to apply the same pathway for various purposes by varying the timing of the extracellular input.

## AT THE FLIP OF A GENE

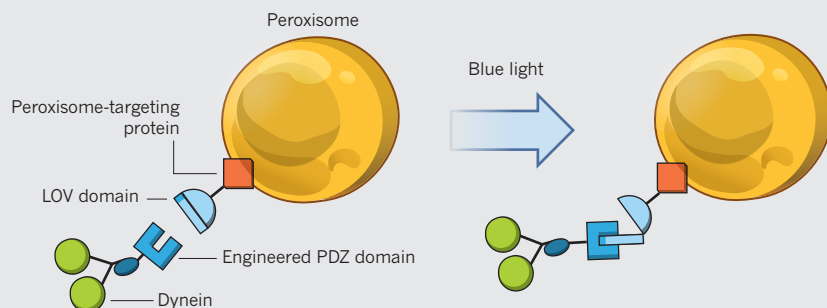
Cell-signalling systems such as Toettcher's affect the activation of genes only after a cascade of intermediate reaction steps. But optogenetic tools can also modify gene expression directly or even induce permanent changes to the genome. For example, Gardner and his colleague Laura Motta-Mena, a biochemist and cell biologist at the University of Texas Southwestern Medical Center in Dallas, have borrowed a light-activated transcription factor from bacteria to activate genes in a range of organisms<sup>8</sup>. At the University of Tokyo, meanwhile, chemist Moritoshi Sato and his colleagues have devised systems that use light to activate CRISPR–Cas9-based gene targeting to achieve high-precision control of gene editing or expression<sup>9,10</sup>.

Optogenetic CRISPR tools such as these will be particularly useful for scientists who want to follow cell behaviour in entire organisms, Hahn says. For example, researchers might want to test whether a cell migrates from one organ to another. With light and CRISPR editing, they could mark the cells that they are interested in with an extra gene encoding something obvious such as green fluorescent protein. Then they could use a microscope to check where those cells go. Sato has speculated that scientists could use the optogenetics–CRISPR combination to investigate how a sequence of mutations turns a cell cancerous or how gene activation in different parts of the brain affects the organ's function.

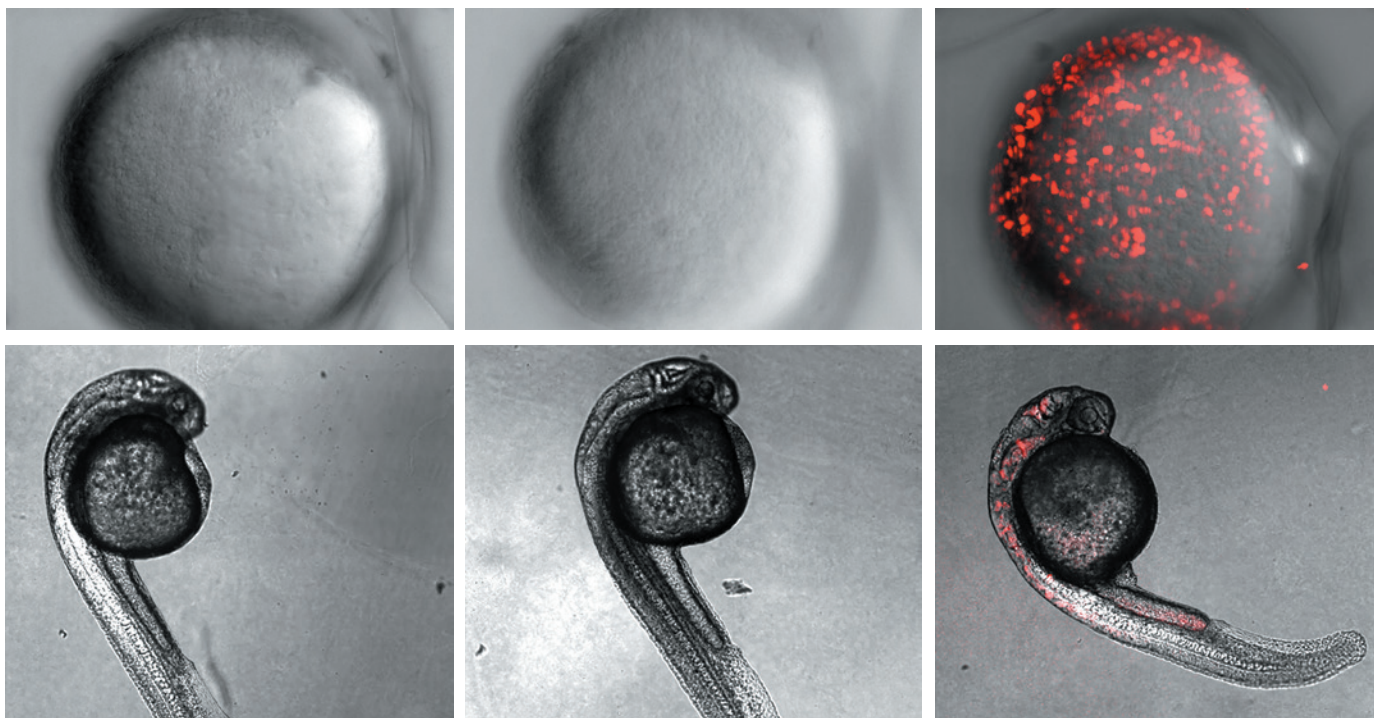
Optogenetics techniques now allow scientists to activate individual genes or proteins with the flick of a light switch, but the next step will be to control multiple processes with a whole spectrum of light. Different proteins could be

## LIGHT SWITCH

To move an organelle, in this case a peroxisome, to the centre of a cell, scientists first attach a light-oxygen-voltage-sensing (LOV) domain to a protein that targets peroxisomes. They also attach an engineered PDZ domain to the motor protein dynein. In the presence of blue light, the LOV opens, the PDZ grabs it and the dynein starts dragging the whole assembly inwards.







NATURE CHEM. BIOL.

Eight-hour- (top) and one-day-old (bottom) zebrafish embryos into which a gene for red fluorescent protein has been added with a sequence that turns on the gene when it is bound by a light-induced transcriptional activator (right). In the dark (left) or when the activator is absent (middle), the gene is switched off<sup>8</sup>.

made sensitive to different colours, so researchers could, for example, flick a blue-light switch to turn on one protein and then a green-light switch to activate another in the same cell. “I would love to see, not one red and one blue, but something like one of those big, old-fashioned organs where you have all kinds of switches and levers,” says Gardner. Unfortunately, at this point, most optogenetic switches play the same note — they usually react to just blue light (although some, such as PhyB, respond to red). But researchers are working on systems that would respond to more diverse colours. Some are even exploring other parts of the electromagnetic spectrum, hoping to activate proteins with microwaves, magnetic fields or radio waves, although that work is in the early stages.

There are other disadvantages of the current optogenetics toolkit. For one, many systems are a bit ‘leaky’ in that they allow some activity even in the dark. And light itself can affect cellular activities such as transcription and signal transduction, points out Masa Yazawa, a stem-cell biologist at the Columbia University Medical Center in New York. This means that scientists should be careful about their negative controls, he says. Just leaving cells in the dark isn’t good enough; rather, scientists should engineer a light-insensitive version of their optogenetic proteins and shine the light on those cells, too.

Another disadvantage is that some light-sensitive systems require a chemical called a chromophore, which scientists have to add if the cells they want to study do not manufacture it. This can be an inconvenience, but it also makes it easy to perform a negative control experiment — the chromophore can simply be left out.

Illumination can also be toxic in large doses. For experiments in which a quick flip of the light switch is all that is needed, this is no big deal. For Kapitein, it takes only a couple of milliseconds to activate the LOV domain and send organelles on their way, so cells have no problem. By contrast, Yazawa wants to grow cells with light-activated genes for days or weeks as they change from stem cells into brain or heart cells. Light toxicity could be a major issue, but fortunately Yazawa’s switches — which are also borrowed from *Arabidopsis* — stay on after they have been stimulated, so he does not have to keep them under constant light<sup>11</sup>. Other scientists, such as Gardner, strobe the light on and off to limit cell exposure while keeping their optogenetic tools activated.

A further problem is that because the tools are so new, they can still be finicky to use. “There’s no plug ‘n’ play,” says Tucker. Every cell biologist with a plan to use light will have to optimize their system, figuring out which optogenetics tools work best for them and identifying the best expression level for their light-sensitive genes. Tucker points out one rookie mistake: using the white light of the microscope to focus samples. White light contains all colours, and will activate the optogenetic sensors. It is better to use filtered light in a colour that will not stimulate the proteins.

Scientists expect that in the future, it will be easier for cell biologists to adopt optogenetics. Researchers are starting to compare different

light-sensitive proteins side-by-side, and their data will help others to select the best tools for their questions. In June 2015, Gardner and Motta-Mena founded a company called Optologix, based in Dallas and New York, to offer standardized kits. Their first product will include the LOV-based gene-activation system they invented, along with an LED lamp.

On the bright side, the lighting part of the package is easy. A bench-top light may do, or the filters and lasers on many microscopes can activate proteins as desired. That accessibility could make light-based tools as standard as microscopes and pipettes in cell biology. “Ten years from now, these will be workhorse tools for everybody in developmental and cell biology,” Toettcher predicts. ■

Amber Dance is a writer in Los Angeles.

1. van Bergeijk, P., Adrian, M., Hoogenraad, C. C. & Kapitein, L. C. *Nature* **518**, 111–114 (2015).
2. Korolchuk, V. I. et al. *Nature Cell Biol.* **13**, 453–460 (2011).
3. Courchet, J. et al. *Cell* **153**, 1510–1525 (2013).
4. Spillane, M., Ketschek, A., Merianda, T. T., Twiss, J. L. & Gallo, G. *Cell Rep.* **5**, 1564–1575 (2013).
5. Strickland, D. et al. *Nature Methods* **9**, 379–384 (2012).
6. Levskaya, A., Weiner, O. D., Lim, W. A. & Voigt, C. A. *Nature* **461**, 997–1001 (2009).
7. Toettcher, J. E., Weiner, O. D. & Lim, W. A. *Cell* **155**, 1422–1434 (2013).
8. Motta-Mena, L. B. et al. *Nature Chem. Biol.* **10**, 196–202 (2014).
9. Nihongaki, Y., Yamamoto, S., Kawano, F., Suzuki, H. & Sato, M. *Chem. Biol.* **22**, 169–174 (2015).
10. Nihongaki, Y., Kawano, F., Nakajima, T. & Sato, M. *Nature Biotechnol.* **33**, 755–760 (2015).
11. Yazawa, M., Sadaghiani, A. M., Hsueh, B. & Dolmetsch, R. E. *Nature Biotechnol.* **27**, 941–945 (2009).



# CAREERS

**TRADE TALK** Venturing from medicine to finance **p.61**

**ACADEMIA** New faculty must embrace work outside the lab [go.nature.com/rq6amb](http://go.nature.com/rq6amb)

**NATUREJOBS** For the latest career listings and advice [www.naturejobs.com](http://www.naturejobs.com)

SIRACHAI ARUNRUGSTICHAU/GETTY



A research diver collects data in a coral reef.

## MARINE BIOLOGY

# Charting sea life

*You want a career in marine biology but your maths is weak. Relax, the basic skills can be mastered.*

BY CHRIS WOOLSTON

Biology attracts all sorts, from number crunchers to big-picture dreamers. These days in science, there's no escape from maths in any scientific discipline, even in one like marine biology, historically lighter on sums than, say, molecular biology or quantitative genetics. But nobody should let maths jitters deter them if their call is to study ocean life.

Although marine biology is built on a foundation of numbers — from the concentration of pores on a shark's snout to the survival rates of seal pups or worm distribution in sea-floor sediments — not every successful marine biologist is a whiz with numbers. Milton Love at the Marine Science Institute at the University of California (UC), Santa Barbara, readily acknowledges that maths is his biggest weakness. "I failed eighth-grade algebra," he says. "And I failed calculus as an undergraduate at UC San Diego. There was a point where I thought I'd have to take calculus 800 times to finally pass." He ended up squeaking through a lower-level calculus course, and went on to build a fulfilling career in research without ever feeling comfortable with the numerical side of his work. "I always managed to finesse the whole thing," he says.

Like many other scientists who struggle with a particular aspect of their research, he simply refused to let a deficiency derail his ambition — an ambition that he had harboured from childhood. "Nobody ever told me that I couldn't be a scientist because I was bad at math," he says. "I just bullied ahead. I was driven." In Love's — and other researchers' — opinion, almost anyone who is truly committed to science can find a niche, even if maths feels like a foreign language. Marine scientists for whom maths is not a strong point need a mix of determination and collaboration to go with their calculations — and the willingness to read a few books, download a video or two and maybe take an online maths and statistics course.

## RUBBISH AT MATHS

Tammy Horton, a marine biologist at the National Oceanography Centre in Southampton, UK, often shares a not-so-secret confession with her students. "I'm very honest," she says. "I say I'm rubbish at maths. A lot of them breathe a sigh of relief." As it happens, Horton's speciality, the taxonomy of small deep-sea crustaceans, does not require ►



► much quantitative skill. To sort out one species from another, she often measures limb lengths or counts hairs, but that is a long way from differential calculus. It is also a long way from the types of multivariate analyses that ecologists, for example, face routinely. “I’m very lucky that I don’t have to use much maths,” she says. “A lot of marine biologists use a huge amount of maths, and it’s getting more mathematical all the time.”

Horton stresses that she did not get into taxonomy because she was trying to avoid hard-core statistics. Instead, she ended up diving into the tiniest details of already tiny creatures because that was what she really wanted to do — study the diversity and adaptations of deep-sea denizens at a very fundamental level. Career paths, she says, should be based on strengths, not on weaknesses. “You shouldn’t choose a career because you have anxiety about statistics,” she says. In her experience, determination can overcome most deficiencies. “The best thing to do is to recognize that maths doesn’t come easily for you,” she says. Armed with that self-awareness, she says, it’s possible to learn skills, erase deficits and find a place in science.

***“I would blame math anxiety more on educational history and less on innate abilities.”***

### PACKAGE DEALS

Kathy Conlan, who researches marine life in the Arctic and Antarctic at the Canadian Museum of Nature in Ottawa, also feels disadvantaged when it comes to maths. “It doesn’t come easy for me,” she says. She is not above asking other people for help with statistics or programming, but she often just ploughs ahead on her own. That is partly because she works at a small institution with fewer options for collaboration, but also because she thinks it is better to “face the hurdles head on”. Before using the statistical package PERMANOVA for analysis of multiple variables in a recent paper, she took a university course on the programme. Even then, she says, analysing the data was somewhat of a struggle. “I was reading and rereading the manual,” she says. “I had to go back and look at university statistics books.”

Conlan’s maths issues go back to her master’s degree, when she had to use punch cards to program the university’s computer (it was the 1970s). She says that young researchers today probably face greater expectations when it comes to mathematical ability. But they do have more resources, including online maths and statistics courses to make up the gaps (see ‘Resources for mathophobes’). “There are so many more ways now to help yourself,” she says.

Statistics programs such as PERMANOVA, and the increasingly popular R, have levelled the playing field, says Steve Haddock, a marine

biologist at the Monterey Bay Aquarium Research Institute in Moss Landing, California, and co-author on a book on computing for biologists. “You don’t have to type in all of the equations, and you don’t have to do the math yourself,” he says. But he warns that canned programs also open the door for big mistakes if users are not thinking carefully about their data — detritus in, detritus out. Scientists who do not feel comfortable with numbers need at least to develop an intuitive sense of the problem that they are trying to address, he says, so that they know which part of the program to use. And, he adds, they need to have at least a general feeling for the data so that they can sort out the plausible results from the outlandish. “If you can’t do all of the calculations, you should at least be able to make a ballpark estimate,” he says.

### MISGUIDED FEAR

Haddock, who used to program his own simple adventure games when he was a kid, says that his proficiency with computers has been a big asset in his career. But he knows that many early-career researchers are not so well prepared. In his view, anxiety about maths and computing should not keep anyone from pursuing science. “It’s easy to think, ‘Other people are better at this than I am,’” he says, “but these things can be overcome.” Besides, he says, fears about maths are often misguided. “It makes me sad to think about people who tell themselves that they’re not good at math,” he says. He believes that many junior scientists who feel that they have a maths deficiency could become fluent with the right encouragement and practical instruction.

From his own experience and conversations with other scientists, Haddock believes that many biologists get counterproductive instruction that erodes their confidence with

numbers. “I would blame math anxiety more on their educational history and less on their innate abilities,” he says. He recalls, for example, a poorly run biology statistics class in his graduate programme. Instead of introducing the students to the stats that they might need to describe their data, the instructor started by mathematically deriving the rationale for the *t*-test — the classic statistical method for determining whether two sets of data differ significantly from one another — which they were unlikely to understand and even less likely to use in the future. Similarly, he believes that many programming classes for scientists dwell on esoteric computing topics instead of on skills that researchers need, such as writing and debugging code to sort through large data sets.

Love says that many of the required or core maths courses for both undergraduate and graduate students seem designed more to weed out degree candidates or to complete a rite of passage than to prepare students for scientific careers. “The first couple of years as a biology major has nothing to do with a career in biology,” he says. “If you can survive the first couple of years, you can find out what biology is all about. It’s not about calculus and physics.” When students come to him with concerns about maths or other parts of their education, he encourages them to look at the big picture. “If they say they like algae, I tell them to hang in there long enough to take an actual algae class,” he says.

### STRENGTH IN NUMBERS

Many researchers have found that a little outside assistance can go a long way when facing mathematical obstacles. Horton says that she sought advice from members of the statistics department when she was getting her PhD, and she still depends on collaboration today. There is a particular statistics-minded



Marine biologist Milton Love researches fish from a manned submersible.

LINDA SNOOK



## HELPING HAND

## Resources for mathophobes

Junior researchers who feel intimidated by maths and computing have options for sharpening their skills and building their confidence, says Steve Haddock, a marine biologist at the Monterey Bay Aquarium Research Institute in Moss Landing, California. His book *Practical Computing for Biologists* (Sinauer, 2011) was written to cut through the confusing clutter of many computing classes and focuses on the 10% of programming techniques that scientists use 90% of the time.

Unfortunately, Haddock says, few books take a similar approach to statistics. But researchers can get a better grasp of their data, including the ability to spot patterns and outliers, with the help of books on data visualization, such as William Cleveland's classic, *The Elements of Graphing Data* (Hobart, 1994), and Stephen Few's *Now You See It* (Analytics, 2009).

Haddock notes that the website Stack Exchange ([stackexchange.com](http://stackexchange.com)) can be a good place to get answers to specific maths and computing questions,

although the online maths community might feel intimidating to someone who is already anxious about the topic. Help is also available on YouTube, including a channel by mathematician Vi Hart ([go.nature.com/mcqc3r](http://go.nature.com/mcqc3r)) that uses art to demonstrate mathematical concepts. The videos help viewers to feel more familiar and comfortable with the numbers around them.

As part of its Elementary Maths for Biologists course, the University of Cambridge, UK, offers free online audio-visual tutorials on a wide range of mathematical topics including scientific notation, logarithms and exponential equations ([go.nature.com/nx8xga](http://go.nature.com/nx8xga)).

For sheer inspiration, scientists who feel inadequate with numbers could read 'Great scientist ≠ good at math', a 2013 *Wall Street Journal* essay by the eminent entomologist E. O. Wilson ([go.nature.com/573cmo](http://go.nature.com/573cmo)). One of his observations: "Many of the most successful scientists in the world today are mathematically no more than semiliterate." **C.W.**

## TRADE TALK

### Venture capitalist



*Bali Muralidhar is an investment manager at London-based MVM Life Science Partners, an international life-sciences growth-capital fund. He describes his move from medicine into business.*

#### Did you always want to work in investment?

No. Initially I wanted to go into academic medicine. I did preclinical training and worked as a junior doctor for a few years before I earned a PhD in translational cancer research from the University of Cambridge, UK. After that, I took a job as a health-care consultant at a small consulting firm and then moved into venture capital.

#### How did your scientific background prepare you for working in venture capital?

Medicine teaches you how to handle pressure. I learned how to deal with different types of people across the spectrum of wealth and personality. Equally, it puts a lot of the decisions I make in business into perspective: although everything is important, they are not life-or-death decisions. My PhD taught me to apply analytical rigour when I am collecting data, planning experiments and thinking things through.

#### What does your role entail?

We look for and invest in early-stage health-care and bioscience companies that we think have the best chance of success in drug or medical-device development and in their potential to help patients. Assessing the science is critical, but of equal importance is assessing whether a product will be clinically and commercially successful.

#### How do you determine if an idea will work?

You have to ask yourself where this product or treatment would fit in to a real-life medical setting. Is it viable in terms of time and resources in the treatment area? Is it going to make the company profitable? These are all things you wouldn't think about when you are doing in-depth research on a particular molecule, say. Just because you find something novel and get onto the front cover of *Nature* does not actually mean that this is going to change people's lives. ■

#### INTERVIEW BY JULIE GOULD

This interview has been edited for length and clarity; see [go.nature.com/nvh18c](http://go.nature.com/nvh18c) for more.

person in her department who is always glad to answer a question or give her much-needed feedback. "People are going to be willing to help you, and they'll do it for free," she says.

Like Horton, Love counts on outside support for his maths. Most of his recent grants have included earmarks — typically about US\$10,000 at a time — specifically for statistical help. As he explains, his main areas of research, such as measuring fish productivity around oil platforms, require a level of analysis that is beyond his reach. He uses that grant money to rent the brainpower of people such as his colleague Mary Nishimoto at the UC Santa Barbara Marine Science Institute or Jeremy Claisse, a biologist at Occidental College in Los Angeles, California. "If I could actually do the statistics myself, which I can't, it would be more efficient," Love says. "But because I'm not doing it, I can do other stuff."

Still, as the importance of maths continues to grow — especially in big-data areas such as ecology, genomics and molecular biology — a little self-sufficiency can go a long way, says Elena Sarropoulou, a marine biologist at the Hellenic Centre for Marine Research in Crete, Greece. "I tell all undergrads and grad students to take a statistics class and to learn the programming language Python," she says. "Just the basics, in order not to be addicted to your

bioinformatician in your lab." She maintains that marine biologists do not have to aspire to mathematical greatness. But they do need to know enough to be able to design an experiment with the appropriate sample size and other parameters to address the problem that they are trying to solve. In her own speciality of molecular biology, she says that researchers too often fail to consider statistical analysis when designing an experiment. She adds that relying solely on a maths expert to interpret results can be risky, because maths-minded people do not always see "the biology behind the data". A biologist with sufficient maths skills will be in the best position to see experiments through from beginning to end, she says.

Until the perfect scientist is invented, every researcher will have some gap in their skill set. The key, Love says, is to find ways to compensate, collaborate and, ultimately, to persevere. Not every researcher can, or even should, hire someone else to do their statistical analyses, he says, but they can find a way to match their aptitudes to their careers and vice versa. "Unless you're just not cut out for the academic life, keep going," he says. "Almost anyone can become a researcher. It's not magic." ■

**Chris Woolston** is a freelance writer in Billings, Montana.

# ONE OF ME

*Unintended consequences.*

BY TARYN HEINTZ

To my darling Bryan,  
At the time, I hated the words ‘difficulty conceiving’. You did too, Bryan, remember? We used to laugh and beg people to be honest, use honest language. It was OK to say Jenny and Bryan are barren or infertile — not reproductively challenged or something similar. But, the language didn’t matter, in the end, we simply could not have children and we were devastated no matter how many different ways people said it.

Remember the day the pamphlet came through the post? I suspect our details were passed to Life Solutions from one of the various clinics we were registered at in our city. It said: “Have a child just like you.” Literally, just like me. A clone baby. What a brilliant idea. People would pay £100,000 and have the option of cloning themselves or their spouse. They could be parents.

“Who should we choose, Bry? A little me or a little you?” I asked.

We chose to clone me because of your weak immune system. That, and we always fancied having a daughter. We wanted to raise a little girl, we wanted to be parents.

The day she was handed over to us, we decided to call her Lila after your grandmother. She didn’t feel like a clone, she didn’t look like a clone, because at the time she looked nothing like me. I was 42, after all. And during her childhood it was fun watching her grow, steering her away from my mistakes.

We encouraged her towards subjects for which we knew she would excel because of her natural aptitude, sports where she would feel a comfortable ease and foods we knew wouldn’t upset her stomach and would keep her fit and lean. Made sure to lather her with SPF and expensive moisturizers.

Lila was like me, only better, healthier. She wouldn’t have my age spots, or struggle to find her place in the world. We could guide her, we had the



inside knowledge. It all seemed so perfect. So easy. She was happy, she was gorgeous.

The day Lila turned 18 you looked at her and said: “God, Jenny, she is so like you, isn’t she?” She was, but her hair was shinier, her clothes hung perfectly on her svelte youthful figure. Did I hear lust in your voice? I thought I did. Even if it wasn’t there. That’s the problem with the seed of doubt, once it’s planted, it’s all you think about.

When Lila was 21, you took her out for drinks. I stayed at home. I couldn’t bear your closeness. I couldn’t bear seeing a younger version of me out with my husband, bonding, joking, while I was fading, ageing.

That’s why people shouldn’t clone themselves, Bryan, because they see the past and the future at the same time. I saw us when I

looked at you and Lila, the old us.

I guess it was my own fault that you enjoyed Lila’s company more than mine. It depressed me to be around her and I couldn’t tell you why. I just sulked.

Was I a mother jealous of her daughter? Or a woman jealous of herself?

How sick of me to accuse you of being attracted to your own daughter, but how can I not want you to be? She’s me. You were attracted to me.

The implications were starting to show. Other clone families were writing to each other, other *mothers* and *fathers* experiencing the same problems. You insisted that you were Lila’s father.

“The End,” you said after we discussed it for the last time. “I’m her father and I love her like a daughter. So she looks like you, so what?! You’re my wife. She’s my daughter. The End.”

I knew better than to bring it up again, but felt a strange sense of reassurance when Life Solutions went bust.

Still, I couldn’t let go. I couldn’t be around Lila without hating her, without resenting her. She is the Jenny I wished I had the discipline and youth to be.

You see, that’s why I withdrew. That’s why I moved out. Seeing her hurts me.

But it’s not enough to simply be away, I need to go away for good. I’m slowly realizing in my quiet solitude that there is only supposed to be one of me in this world. One Jenny for one Bryan, and you deserve the better version.

This isn’t a love note, this is a goodbye letter. I know it’s clichéd and cowardly, but I’ve taken all the sleeping pills in the bottle Dr Greenfield gave me. Every. Last. One. Washed them down with cold vodka and now I await sleep. It’s not a spectacular way to go, but I guess I had *difficulty conceiving* a better way.

I love you,  
Jenny. ■

**Taryn Heintz** currently lives in Cardiff, Wales, and is from Seattle, Washington. She likes to write in her spare time and enjoys learning new languages.

ILLUSTRATION BY JACEY